Central limit theorems with a rate of convergence for dynamical systems

Olli Hella

March 6, 2020

1 Abstract

Central limit theorems are some of the most classical theorems in the theory of probability. They have also been actively studied in the field of dynamical systems. In the first article of this thesis an adaptation of Stein's method, introduced by Charles Stein in 1972 is presented. Our adaptation gives new correlation-decay conditions for both univariate and multivariate observables under which central limit theorem holds for time-independent dynamical systems. When these conditions are satisfied, this adaptation also yields estimates for convergence rates. We also present a scheme for checking these conditions and consider it in two example models.

In the second article the scope of this adaptation is extended further to time-dependent dynamical systems. The applicability of this method is shown for time-dependent expanding circle maps and also for quasistatic dynamical systems, which is a new research area introduced recently by Dobbs and Stenlund.

The third article considers time-dependent compositions of Pomeau-Manneville-type intermittent maps. For this model we also establish central limit theorems with a rate of convergence. This article uses the results in the second article and earlier work of Juho Leppänen on the functional correlation bounds for Pomeau-Manneville maps with time-dependent parameters. Quasistatic systems are also further studied and we present general conditions under which a multivariate CLT for quasistatic systems holds.

In the fourth article we study random compositions of transformations. We prove a theorem on almost sure convergence of the variance of normalized and fiberwise centered Birkhoff sums. This in combination with earlier results, such as the theorems in the second article, can be used to establish quenched central limit theorems with a rate of convergence for random dynamical systems. Two examples which use the theorem in the fourth article are provided in the second and third article.

2 Acknowledgements

First of all, I want to thank Mikko Stenlund for being my advisor through this long PhD project. Mikko's advice has helped me enormously to become a better researcher and to improve my mathematical thinking. He has also had very good ideas for research projects and is able to put difficult concepts into words in a very clear precise manner.

I am also very thankful for Antti Kupiainen for the opportunity to do this PhD thesis in the mathematical physics group and also for being the custos of the thesis defence. I thank Peter Balint for being my opponent in the thesis defence and also for his hospitality during my visit in Budapest. I also thank Esa Järvenpää and Antti Käenmäki for pre-examining the thesis.

I thank Juho Leppänen for being nice co-worker. It has been pleasant to work with him and we have had good discussions which have deepened my understanding of our shared research projects and also given me more motivation.

I had three year break from university after having my Master's degree. During that time I tried to get funding for a PhD project without luck. I am very grateful to Jouni Luukkainen who helped and gave me encouragement back then by giving feedback in the process of writing articles about ideas in my Master's thesis. I'd also like to thank all my teachers from elementary school to university who have supported my dreams and given me inspiration. I also thank Kalle Kytölä whose long reply to my email, where I asked about the possibilities of doing a PhD, finally led me to find a advisor and funding for writing this thesis. I have received funding for this research from Academy of Finland, Jane and Aatos Erkko Foundation, and Emil Aaltosen Säätiö to whom I am very grateful.

My greatest weakness as a researcher[1] is that I like leisure time and hanging out with friends. I thank all my friends, especially Perttu Kiviharju and Sofia Holopainen for occasional lunches which have cheered me up during the work day. I also thank Joona Sahamies, Mira & Jani Heino and Minna-Johanna Lauri-Haikala & Topias Haikala who let me live in their homes for some time during a pipe renovation in my apartment.

Lastly, I want to thank my mother and father for always being very supportive to me and my brother Ossi, and sister Anni - with her husband Lauri - with whom we share a same Hellaish type of humour. It is always fun to meet them.

[1] beside being interested in almost every subject, even economy, church history and semantics nowdays

3 List of included articles

This thesis consists of an introduction and four research articles. In the introduction these research articles are referred with letters **(A)**–**(D)**

(A) Olli Hella, Juho Leppänen and Mikko Stenlund. Stein's method of normal approximation for dynamical systems. To appear in *Stochastic and Dynamics.*

(B) Olli Hella. Central limit theorems with a rate of convergence for sequences of transformations. arXiv:1811.06062

(C) Olli Hella and Juho Leppänen. Central limit theorems with a rate of convergence for time-dependent intermittent maps. To appear in *Stochastic and Dynamics.*

(D) Olli Hella and Mikko Stenlund. Quenched normal approximation for random sequences of transformations. *Journal of Statistical Physics, 178(1), 1-37.*

Article **(A)** is a joint work with Juho Leppänen and my advisor Mikko Stenlund. The initial idea to study the applicability of Stein's method for dynamical systems came from Stenlund. Stenlund did the introductory part, and abstract scheme and dispersing billiard model parts in section 7 of the paper. In Section 4 I did the part concerning the estimates for fourth order term while Leppänen did the estimates for third order term. For other parts of the paper we contributed an equal amount with Leppänen.

Article **(C)** is a joint work with Leppänen. We gave an equal amount of contribution to the paper. I concentrated to the parts requiring knowledge on Stein's method for the time-dependent maps I studied in paper **(B)** and random dynamical systems studied in **(D)**, while Leppänen concentrated on the parts on the intermittent maps and functional correlation bounds.

Article **(D)** is a joint work with Stenlund. Stenlund planned the overall strategy of the paper and did most of the intro and appendices. He also wrote the random dynamical systems theory parts. I did the proofs and estimates in sections 2–4 and parts of the appendices and intro.

CENTRAL LIMIT THEOREMS WITH A RATE OF CONVERGENCE FOR DYNAMICAL SYSTEMS

1. Preliminaries

1.1. **GOAL.** The focus in the articles of this thesis are central limit theorems with a rate of convergence. Beyond convergence rates, we also give some results where a concrete number for the upper bound of the difference of the law of a suitably normalized and centered Birkhoff sum and the normal distribution can be computed.

The articles in this dissertation are

(A) Stein's method of normal approximation for dynamical systems

(B) Central limit theorems with a rate of convergence for sequences of transformations

(C) Central limit theorems with a rate of convergence for time-dependent intermittent maps

(D) Quenched normal approximation for random sequences of transformations

1.2. **Notations and conventions.** In this introduction we denote $\mathbb{N} = \{1, 2, 3, ...\}$ and $\mathbb{N}_0 = \{0, 1, 2, ...\}$. Let $T\colon X \to X$ be a map. Then $T^0\colon X \to X : x \mapsto x$, i.e., T^0 is an identity map denoted by Id. Similarly, if $(T^n)_{n=1}^{\infty}$ is a sequence of transformations from X to X, then $T_i \circ ... \circ T_1 = \mathrm{Id}$ for $i = 0$. For sums $\sum_{i=j}^{k}$, where $k < j$, we define the sum to be 0. We also use the convention $0^0 = 1$. The symbol Z stands for a random variable with standard normal distribution $\mathcal{N}(0, 1)$ unless otherwise stated.

1.3. **What is a dynamical system?** The term *dynamical system* has some ambiguity. We follow the book [39], where a dynamical system is described as consisting of the following three components:

i) *State (or phase) space.* This consists of points that represent the possible states of the system.

ii) *Time.* Time is either continuous or discrete and it can extend both to the future and the past or only to the future.

iii) *The time evolution law.* This is the rule that determines how the state of the system evolves in time.

In this introduction we only consider dynamical systems with discrete time extending only to the future, i.e., time is modelled by $\mathbb{N}_0$. In article **(A)** there is one result that applies to continuous time, but we omit it here, since it is not a key theorem in **(A)**.

We denote the state space of the system by X. Four types of dynamical systems are considered in the articles of this thesis. The first type is a dynamical system consisting of a measure space $(X, \mathcal{B}, \mu)$ combined with a measure preserving transformation $T\colon X \to X$. Being measure-preserving means that T is measurable and $\mu(T^{-1}A) = \mu(A)$ for every $A \in \mathcal{B}$. In this type of dynamical system if the state of the system at the moment 0 is x, then at the time $t \in \mathbb{N}_0$ it is $T^t(x)$, where T^t means that transformation T is applied t times. Dynamical system with one measure-preserving map are also stationary, i.e.,

$\mu(A_0 \cap T^{-1}A_1 \cap \cdots \cap T^{-n}A_n) = \mu(T^{-1}A_0 \cap T^{-2}A_1 \cap \cdots \cap T^{-n-1}A_n)$ for every $n \in \mathbb{N}_0$ and $A_0, ..., A_n \in \mathcal{B}$. Therefore we abbreviate this type of system by SDS, where first S is from the word stationary.

The second type, a time-dependent dynamical system (TDDS), is similar to the first one, except now instead of one transformation T, the time evolution is defined by sequence of transformations $T_1, T_2, ...$, where $T_i \colon X \to X$ for all $i \in \mathbb{N}$. Now if the state of the system at time 0 is x, then at time $t \in \mathbb{N}_0$ it is $T_t \circ T_{t-1} \circ \cdots \circ T_1(x)$.

The third type under investigation are quasistatic dynamical systems (QDS). This type of dynamical system was introduced in [19] quite recently. It models a scenario where the time-evolution law of a physical system changes infinitesimally slowly under external influence on the system. The details of the definition of a QDS are given later in Section 6.

The fourth type of dynamical systems are random dynamical systems (RDS) in which the law of time-evolution is picked randomly. The definition of RDS is given in Section 7.

The type SDS is studied in article **(A)**. A result for continuous time version of measure preserving transformation is also given in **(A)**. Article **(B)** focuses on TDDS although example models for QDS and RDS are also studied. In article **(C)** we study special cases of TDDS, RDS and QDS, where transformations are Pomeau-Manneville maps. Furthermore a general result for proving QDS central limit theorems (CLT) is given. Article **(D)** is all about proving quenched limit theorems for RDS. The RDS parts of articles **(B)** and **(C)** also apply the results of **(D)**. The results in each paper of this thesis are concentrated around CLTs.

1.4. **Outline of the introduction.** In Section 2 we discuss briefly dynamical systems in general and give a short introduction to CLTs. In Section 3 we introduce the method that we use to prove CLTs in our research articles. In Sections 4, 5, 6 and 7 we give rigorous definitions of these systems, present the main results of **(A)**-**(D)** and discuss how these results are linked to earlier research on these systems.

2. Central limit theorems

Nature is full of systems obeying laws of physics. Building scientific theories is based on measuring those systems. We call the result of those measurements observations. The observations might be about temperature, length, mass, energy, velocity etc. The measured physical quantity itself is called an observable. In the end the observations usually consist of numerical value(s) and a unit of measurement, for example 4.25 kg. The practical value of science is the ability to make predictions about the future values of observations given some values of previous measurements. In an ideal case the future values of some observations can be predicted with very small error margins for very long time periods as in the case of predicting the future locations of the planets in the solar system. In some other cases like weather forecasting, the system in question, i.e., the atmosphere of the earth, is chaotic, and therefore the predictions lose accuracy fast. In those cases we might still be able to make some statistical predictions about future observations, for example their long time averages. In this thesis observations are mathematical quantities in $\mathbb{R}$ or $\mathbb{R}^d$, which are values of an observable, which itself is a function from the state space X of the system to $\mathbb{R}$ or $\mathbb{R}^d$, where $d \in \mathbb{N}$.

For the types of dynamical systems introduced in the previous section, RDS is the only one with unpredictable time-evolution. One might wonder how the other three types of systems produce any statistical behaviour, them being completely deterministic. This is achieved by modeling the initial state of the system X by a probability measure μ on X. A physical interpretation of this construction is that our knowledge of the state of the system X is imprecise or incomplete so we can model this lack of knowledge by a probability distribution. For SDS, TDDS and QDS this probability distribution then evolves deterministically in time, so all indeterminacy is hidden in the uncertainty of the initial state of the system.

Consider for example a measure-preserving dynamical system $(X, \mathcal{B}, \mu, T)$. Let $f\colon X \to \mathbb{R}^d$ be an observable that we are interested in. We can ask for example what is the time average $n^{-1}\sum_{t=0}^{n-1} f\circ T^t$ of the first n observations of observable f. A famous theorem by Birkhoff states that when μ is ergodic[1] and $f \in L^1(X, \mathcal{B}, \mu)$, then

$$\lim_{n\to\infty} n^{-1}\sum_{i=0}^{n-1} f\circ T^i = \int_X f \,\mathrm{d}\mu$$

almost surely. This result resembles the strong law of large numbers. The strong law of large numbers states that for i.i.d. random variables satisfying $\mathbb{E}|X_i| < \infty$ and $\mathbb{E}[X_i] = 0$ it holds that

$$\lim_{n\to\infty} n^{-1}\sum_{i=0}^{n-1} X_i = 0 \qquad \text{almost surely.}$$

This can be shown by using Birkhoff's theorem. The physical interpretation of Birkhoff's theorem is that a time average of an observable equals the corresponding space-avarage almost surely; or in other words if we choose an initial state x randomly with respect to measure μ and let the system evolve through time, then with probability 1 the time average of observables converge to the constant $\int_X f\,\mathrm{d}\mu$.

We can then ask about the distribution of differently scaled sum $\lim_{n\to\infty} n^{-1/2}\sum_{i=0}^{n-1} X_i$. We are interested on the conditions under which random variables X_i satisfy the so called central limit theorem (CLT), namely when does

$$\lim_{n\to\infty} n^{-1/2}\sum_{i=0}^{n-1} X_i \xrightarrow{d} \mathcal{N}(0,\sigma^2)?$$

Or in other words: when does $\lim_{n\to\infty} \mathbb{P}(n^{-1/2}\sum_{i=0}^{n-1} X_i \le x) = \Phi_{\sigma^2}(x)$, $\forall x \in \mathbb{R}$, where Φ_{σ^2} is the cumulative distribution function of the normal distribution with mean 0 and variance σ^2? This question has received a huge amount of interest in the field of probability theory for over two centuries [25]. Letting $X_i = f\circ T^i$, we can ask this same question in the context of dynamical systems. It has also been studied a long time, see for example [67] and [63] for some early results.

For practical applications of the theory of dynamical systems it would be essential to also know how fast the distribution of scaled random variables approach the normal distribution, i.e., to have a rate of convergence to the normal distribution. A way to measure the difference between two distributions is needed to establish a rate of convergence. This can be accomplished in the following way: Let $\mathscr{H}$ be a non-empty class of functions $h\colon \mathbb{R}\to\mathbb{R}$. Let X, Y be two random variables with distributions L_X and L_Y. Define

$$d_{\mathscr{H}}(L_X, L_Y) = \sup_{h\in\mathscr{H}} |\mathbb{E}[h(X)] - \mathbb{E}[h(Y)]|.$$

Now $d_{\mathscr{H}}$ is a (pseudo)metric in the class of probability distributions. We write $d_{\mathscr{H}}(X,Y) = d_{\mathscr{H}}(L_X, L_Y)$.

[1]Measure μ is ergodic if for all $A \in \mathcal{B}$ that satisfy $T^{-1}A = A$, it holds that $\mu(A) \in \{0,1\}$.

Our main contribution to the study of CLTs for dynamical systems is a general adaptation of Stein's method [68] to the DS setup. For Stein's method, applied in **(A)**–**(C)**, there exists a natural choice for $\mathscr{H}$, the class of 1-Lipschitz maps $h\colon \mathbb{R} \to \mathbb{R}$. We write $\mathscr{W} = \{h : |h(x) - h(y)| \leq |x - y|\}$. The Wasserstein distance of two random variables X and Y is now defined as

$$d_{\mathscr{W}}(X, Y) = \sup_{h \in \mathscr{W}} |\mathbb{E}[h(X)] - \mathbb{E}[h(Y)]|.$$

Another commonly used distance called Kolmogorov distance $d_{\mathscr{K}}$ is defined by class $\mathscr{K} = \{1_{(-\infty,x]} \,:\, x \in \mathbb{R}\}$. Kolmogorov distance of two random variables X and Y is $\sup_{t\in\mathbb{R}} |\mathbb{P}(X \leq t) - \mathbb{P}(Y \leq t)|$, the maximum difference of the cumulative distribution functions of X and Y.

Next we look at normal distributions more closely and discuss some ways to characterize them, which can then be used in proofs of CLTs.

2.1. **Normal distribution.** The normal distribution with mean μ and variance σ^2 is denoted by $\mathcal{N}(\mu, \sigma^2)$. We say that a property P characterizes $\mathcal{N}(\mu, \sigma^2)$ if $\mathcal{N}(\mu, \sigma^2)$ is the only distribution that satisfies P. Thus we can prove that some random variable X is normally distributed if one of these properties is satisfied. We introduce some of these, of which the last one is crucial to our articles **(A)**–**(C)**:

Probability density function. Normal distribution $\mathcal{N}(\mu, \sigma^2)$ with $\sigma^2 > 0$ is characterized by its probability density function

$$\phi_{\sigma^2,\mu}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

This also defines the corresponding cumulative distribution function

$$\Phi_{\sigma^2,\mu}(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{x} e^{\frac{(t-\mu)^2}{2\sigma^2}} \, \mathrm{d}t.$$

Thus $Y_n \xrightarrow{d} \mathcal{N}(\mu, \sigma^2)$ if and only if

$$\lim_{n\to\infty} \mathbb{P}(Y_n \leq x) = \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{x} e^{\frac{(t-\mu)^2}{2\sigma^2}} \, \mathrm{d}t.$$

for all $x \in \mathbb{R}$.

Characteristic function. Let X be a random variable. Its characteristic function is defined by

$$\varphi(t) = \mathbb{E}[e^{itX}].$$

Characteristic function determines the probability distribution, i.e., if X and Y are two random variables with the same characteristic function then they also have the same distribution. The normal distribution $\mathcal{N}(\mu, \sigma^2)$ is characterized with the characteristic function $\varphi(t) = e^{i\mu t - \sigma^2 t^2/2}$ [23].

Stein's characterization. Stein found that the normal distribution $\mathcal{N}(0, \sigma^2)$ can be characterized by the following condition (see [16]): Let Y be a random variable. Then Y has a normal distribution $\mathcal{N}(0, \sigma^2)$ if and only if

$$\mathbb{E}[YA(Y)] = \sigma^2 \mathbb{E}[A'(Y)]$$

for all absolutely continuous functions A for which these expectations exist. If then $Y \sim \mathcal{N}(\mu, \sigma^2)$, the above characterization yields $\mathbb{E}[(Y - \mu)A(Y - \mu)] = \sigma^2 \mathbb{E}[A'(Y - \mu)]$.

Next we look at multivariate normal distributions.

2.2. **Multivariate normal distribution.** Let $d \in \mathbb{N}$ and $\Sigma \in \mathbb{R}^{d\times d}$. A random vector $Z = (Z_1, Z_2, ..., Z_d)^T$ has a multivariate normal distribution $\mathcal{N}(\mu, \Sigma)$, if all linear combinations $Y = a_1Z_1 + ... + a_dZ_d$ are normally distributed, $\mathbb{E}[Z] = \mu$ and $\mathbb{E}[(Z-\mu)(Z-\mu)^T] = \Sigma$.

Central limit theorems can be generalized for d-dimensional random vectors. We say that a sequence $(X_i)_i^\infty$ of d-dimensional random vectors defined on the same probability space satisfy a CLT if the scaled sum $n^{-1/2}\sum_{i=0}^{n-1} X_i$ converges in distribution to a multivariate normal distribution.

When Σ is positive-definite, $\mathcal{N}(\mu, \Sigma)$ is characterized by its probability density function

$$\phi_\Sigma(x) = \frac{e^{-\frac{1}{2}(x-\mu)\cdot\Sigma^{-1}(x-\mu)}}{\sqrt{(2\pi)^d \det \Sigma}}$$

and its characteristic function

$$\varphi(t) = e^{it\cdot\mu - \frac{1}{2}t\cdot(\Sigma t)}.$$

As in the univariate case, there also exists a Stein characterization [64] for random vectors Y: Let the matrix $\Sigma \in \mathbb{R}^{d\times d}$ be symmetric and positive definite. Then $Y \sim \mathcal{N}(0, \Sigma)$ if and only if Y satisfies

$$\mathbb{E}[\operatorname{tr}\Sigma D^2 A(Y) - Y \cdot \nabla A(Y)] = 0$$

for all $A \in C^3(\mathbb{R}^d, \mathbb{R})$.

Armed with these concepts, we may now introduce the Stein method which is perhaps the most central theme in this dissertation.

3. Stein's method

3.1. **Short introduction to the method.** Stein introduced a new method for proving central limit theorems in his seminal paper [68]. We now give a short outline how the method works in the settings of **(A)** and **(B)**.

Consider random variables $X_0, X_1, ..., X_{N-1}$ with expectation $\mathbb{E}[X_i] = 0$ for every $i \in \{0, ..., N-1\}$. Assume also that X_i's are weakly dependent in a way described in **(A)** and **(B)**[2]. Denote $W = N^{-1/2}\sum_{i=0}^{N-1} X_i$ and $\sigma^2 = \operatorname{Var}(W)$. Let $h\colon \mathbb{R} \to \mathbb{R}$, and let $\Phi_{\sigma^2}(h) = \frac{1}{\sqrt{2\pi\sigma^2}}\int_{\mathbb{R}^d} e^{-\frac{w^2}{2\sigma^2}} h(w)\,\mathrm{d}w$, i.e., $\Phi_{\sigma^2}(h)$ is the expectation of h with respect to $\mathcal{N}(0, \sigma^2)$. Stein's method gives an upper bound on $|\mathbb{E}[h(W)] - \Phi_{\sigma^2}(h)|$. This is done by solving the so called Stein equation:

$$\sigma^2 A'(w) - wA(w) = h(w) - \Phi_{\sigma^2}(h). \tag{1}$$

It can be shown (see [16] and Lemma 3.2 in **(A)**) that when h is absolutely continuous, the solution A for this equation satisfies

$$\|A\|_\infty \le 2\|h'\|_\infty, \quad \|A'\|_\infty \le \sqrt{2/\pi}\,\sigma^{-1}\|h'\|_\infty \quad \text{and} \quad \|A''\|_\infty \le 2\sigma^{-2}\|h'\|_\infty. \tag{2}$$

Taking the expectation with respect to the distribution of W in (1) yields $|\sigma^2\mathbb{E}[A'(W)] - \mathbb{E}[WA(W)]| = |\mathbb{E}[h(W)] - \Phi_{\sigma^2}(h)|$. When $\|h'\|_\infty \le 1$, we have $|\mathbb{E}[h(W)] - \Phi_{\sigma^2}(h)| \le d_{\mathscr{W}}(W, Z)$, where $Z \sim N(0, \sigma^2)$. Actually

$$\sup_{\|h'\|_\infty \le 1} |\mathbb{E}[h(W)] - \Phi_{\sigma^2}(h)| = d_{\mathscr{W}}(W, Z).$$

[2]Details of weak dependence are not relevant for understanding the basic ideas behind the following outline.

Furthermore assuming $\|h'\|_\infty \le 1$ in (2) yields the bounds $\|A\|_\infty \le 2$, $\|A'\|_\infty \le \sqrt{2/\pi}\,\sigma^{-1}$ and $\|A''\|_\infty \le 2\sigma^{-2}$. Putting the previous observations together gives

$$d_{\mathscr{W}}(W,Z) \le \sup_{\{A:\|A\|_\infty\le 2,\, \|A'\|_\infty \le \sqrt{2/\pi}\,\sigma^{-1},\, \|A''\|_\infty \le 2\sigma^{-2}\}} |\sigma^2\mathbb{E}[A'(W)] - \mathbb{E}[WA(W)]|.$$

We are now left to bound $|\sigma^2\mathbb{E}[A'(W)] - \mathbb{E}[WA(W)]|$. Below we give an outline of an argument which shows that $|\sigma^2\mathbb{E}[A'(W)] - \mathbb{E}[WA(W)]|$ and thus $d_{\mathscr{W}}(W,Z)$ is close to zero with the above assumptions on the random variables X_i:

Let $K, n \le N$ and write $W_n = N^{-1/2}\{\sum_{i=0}^{N-1} X_i - \sum_{i=\max\{0,n-K\}}^{\min\{n+K,N-1\}} X_i\}$, i.e., W_n is the same as W except that some terms near the index n are removed from the sum. Now we may write

$$\mathbb{E}[WA(W)] = \mathbb{E}[N^{-1/2}\sum_{i=0}^{N-1} X_i(A(W) - A(W_i))] + \mathbb{E}[N^{-1/2}\sum_{i=0}^{N-1} X_i A(W_i)]. \tag{3}$$

Note that if the X_i's are independent, then the last term is exactly zero. When the X_i's are only weakly dependent, it seems reasonable that we might find a small upper bound to the last term. In Theorem 2.3 of **(A)**, we just assume a uniform bound on each $|\mathbb{E}[X_iA(W_i)]|$. This bound has to be calculated separately for each dynamical system in question. An abstract scheme of how this might be achieved is introduced on Section 7 of **(A)**.

For the first term on the right side of (3) consider first the case where A' is constant. Then this term is exactly

$$N^{-1}\sum_{i=0}^{N-1}\sum_{j=\max\{0,n-K\}}^{\min\{n+K,N-1\}} \mathbb{E}[X_iA'(W)X_j]. \tag{4}$$

If the pair correlations $\mathbb{E}[X_iX_j]$ tend to zero when $|j-i|$ increases and K is large enough, then $\sum_{i=0}^{N}\sum_{j=\max\{0,n-K\}}^{\min\{n+K,N-1\}} \mathbb{E}[X_iX_j] \approx N\sigma^2$, i.e., (4) $\approx \sigma^2\mathbb{E}[A'(W)]$. If on the other hand A' is not constant but A'' is bounded then the same argument still works (with an added error term). Under these vaguely described conditions for smoothness of A and correlation decay for $X_0, ..., X_{N-1}$ we thus have $\mathbb{E}[WA(W)] \approx \sigma^2\mathbb{E}[A'(W)]$, which in turn implies that $d_{\mathscr{W}}(W,Z) \approx 0$. Rigorous computations to show this are found in **(A)** and **(B)**.

3.2. **Prior studies.** The only appearances of Stein's method in the setting of CLTs in dynamical systems were in [31, 42] for univariate CLT in some special cases, however without deriving estimates of the rate of convergence. Beside these, the method has also been applied in the setup of Poisson limits; see [18, 32, 35, 37, 62]. Article **(A)** is a first general attempt to apply Stein's method to normal approximation for dynamical systems. The results proven in **(A)** are in SDS setting, with both discrete and continuous time and for both univariate and multivariate observables. These results give CLTs with a rate of convergence.

In paper **(B)** results similar to those of **(A)** are proven in the context of TDDS. Actually the main theorems of **(B)** consider sequences $(X_i)_{i=0}^{N-1}$ of random variables that do not have to relate to dynamical systems at all. However, defining $X_i = f \circ T_i \circ ... \circ T_1$, where f is an observable and T_j a transformation for every $j = 1, ..., i$ gives a natural way to apply main theorems of **(B)** to dynamical systems. These main results are also applied to RDS and QDS models in the same paper and to special kind of intermittent TDDS and QDS in article **(C)**.

Stein's method has been researched thoroughly in the theory of probability, where it originated due to Stein himself. It has been applied for example to Poisson [7, 15], exponential [28], binomial [24] and gamma distributions [50]. Articles **(A)** and **(B)** consider Stein's method also on multivariate normal distribution on which there has also been a lot research done in the field of probability theory; see, e.g., [12, 13, 16, 29, 30, 34, 51, 52, 56, 64–66].

4. Stationary dynamical system

In this Section we examine the classical type of dynamical system consisting of a measure space $(X, \mathcal{B}, \mu)$ combined with a measure preserving transformation $T\colon X \to X$. Time is modelled by $\mathbb{N}_0$. As discussed before, this type of system is also stationary.

4.1. Previous research.

A result that gives bounds on the distance of one distribution to a (multivariate) normal distribution is called a normal approximation theorem. These have been obtained for both univariate and multivariate cases in the SDS setting. Some of the most relevant papers on normal approximation and SDS CLTs are introduced below before stating our results.

Liverani [48] proves some CLTs for SDS with proofs based on martingale approximations. Dubois [22] has proved normal approximation theorem for uniformly expanding circle maps with $Cn^{-1/2}$ convergence rate for Kolmogorov distance, where C can be computed. Gouëzel has also proved normal approximation theorem with $O(n^{-1/2})$ convergence rate for Kolmogorov distance, which is applied for Pomeau-Manneville maps [33]. Pène [58] proves CLT for billiard model with $O(n^{-1/2+\epsilon})$, $\epsilon > 0$, convergence rate for Kolmogorov distance. For normal approximation theorems in the multivariate case, see Jan [38] and Pène [59]. A very recent paper [5] by Antoniou and Melbourne obtains *"the first results on convergence rates in the Prokhorov metric for the weak invariance principle (functional central limit theorem) for deterministic dynamical systems"*. These results hold for a large scale of systems, for example dispersing billiards and intermittent maps, which we have also researched. The result that resembles our results in paper **(A)** most closely is found on paper [60] by Pène and will be discussed in the end of this section after presenting our results.

4.2. Main results.

Let $(X, \mathcal{B})$ be a measurable space with initial measure μ. We assume that $T\colon X \to X$ preserves μ, i.e., $\mu(T^{-1}A) = \mu(A)$ for every $A \in \mathcal{B}$. In other words, we work in SDS setup.

We consider a multivariate observable $f\colon X \to \mathbb{R}^d$ with $\mu(f) = 0$. Denote $f^k = f \circ T^k$. The purpose of the next theorem is to give an upper bound for the difference of distribution of $W = W(N) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} f^k$ and $\mathcal{N}(0, \Sigma)$, where $\Sigma = \mu(f \otimes f) + \sum_{n=1}^{\infty} (\mu(f^n \otimes f) + \mu(f \otimes f^n))$ is the limit of the covariance matrices of $W(N)$. This difference is measured in the following way: For every three times differentiable $h\colon \mathbb{R}^d \to \mathbb{R}$ we estimate the difference $\mu(h(W)) - \Phi_\Sigma(h)$, where

$$\Phi_\Sigma(h) = \frac{1}{\sqrt{(2\pi)^d \det \Sigma}} \int_{\mathbb{R}^d} e^{-\frac{1}{2} w \cdot \Sigma^{-1} w} h(w) \, \mathrm{d}w,$$

i.e., the expectation of h with respect to $\mathcal{N}(0, \Sigma)$. Naturally $\mu(h(W)) - \Phi_\Sigma(h)$ is zero, if $W \sim \mathcal{N}(0, \Sigma)$, and close to zero if W is almost normally distributed (and h has moderate second and third order partial derivatives).

First we need to introduce some notations. Let $\alpha = \{1, ..., d\}$. The components of f^i are denoted by $f^i_\alpha = (f \circ T^i)_\alpha$.

Let $N \in \mathbb{N}_0$ and $K \in \mathbb{N}_0 \cap [0, N-1]$. Denote

$$[n]_K = [n]_K(N) = \{k \in \mathbb{N}_0 \cap [0, N-1] \,:\, |k - n| \le K\},$$

and

$$W^n = W - \frac{1}{\sqrt{N}} \sum_{k \in [n]_K} f^k \tag{5}$$

for all $n \in \mathbb{N}_0 \cap [0, N-1]$. In other words, W^n differs from W by a time gap (within $[0, N-1]$) of radius K, centered at time n.

The next theorem gives a bound for $|\mu(h(W)) - \Phi_\Sigma(h)|$ which depends on the choice of K. The best choice of K depends on the dynamical system, but $K \sim C \log N$ works when the correlations described in the conditions decay exponentially (see Corollary 2.2 in **(A)**).

The proof is based on solving Stein's equation $\operatorname{tr} \Sigma D^2 A(w) - w \cdot \nabla A(w) = h(w) - \Phi_\Sigma(h)$ for the multivariate normal distribution $\mathcal{N}(0, \Sigma)$.

Theorem 4.1. *Let $f : X \to \mathbb{R}^d$ be a bounded measurable function with $\mu(f) = 0$. Let $h : \mathbb{R}^d \to \mathbb{R}$ be three times differentiable with $\|D^k h\|_\infty < \infty$ for $1 \le k \le 3$. Fix integers $N > 0$ and $0 \le K < N$. Suppose that the following conditions are satisfied:*

(A1) *There exist constants $C_2 > 0$ and $C_4 > 0$, and a non-increasing function $\rho : \mathbb{N}_0 \to \mathbb{R}_+$ with $\rho(0) = 1$ and $\sum_{i=1}^\infty i\rho(i) < \infty$, such that*

$$\begin{aligned} |\mu(f_\alpha f^k_\beta)| &\le C_2\, \rho(k) \\ |\mu(f_\alpha f^l_\beta f^m_\gamma f^n_\delta)| &\le C_4 \min\{\rho(l), \rho(n-m)\} \\ |\mu(f_\alpha f^l_\beta f^m_\gamma f^n_\delta) - \mu(f_\alpha f^l_\beta)\mu(f^m_\gamma f^n_\delta)| &\le C_4\, \rho(m-l) \end{aligned}$$

hold whenever $k \ge 0$; $0 \le l \le m \le n < N$; $\alpha, \beta, \gamma, \delta \in \{\alpha', \beta'\}$ and $\alpha', \beta' \in \{1, \dots, d\}$.

(A2) *There exists a function $\tilde\rho : \mathbb{N}_0 \to \mathbb{R}_+$ such that*

$$|\mu(f^n \cdot \nabla h(v + W^n t))| \le \tilde\rho(K)$$

holds for all $0 \le n < N$, $0 \le t \le 1$ and $v \in \mathbb{R}^d$.

(A3) *f is not a coboundary in any direction.*

Then

$$\Sigma = \mu(f \otimes f) + \sum_{n=1}^\infty (\mu(f^n \otimes f) + \mu(f \otimes f^n)) \tag{6}$$

is a well-defined, symmetric, positive-definite, $d \times d$ matrix; and

$$|\mu(h(W)) - \Phi_\Sigma(h)| \le C_* \left(\frac{K+1}{\sqrt{N}} + \sum_{i=K+1}^\infty \rho(i) \right) + \sqrt{N}\tilde\rho(K), \tag{7}$$

where

$$C_* = 12 d^3 \max\{C_2, \sqrt{C_4}\} \left(\|D^2 h\|_\infty + \|f\|_\infty \|D^3 h\|_\infty\right) \sum_{i=0}^\infty (i+1)\rho(i) \tag{8}$$

is independent of N and K.

The normal approximation for W is given in the form of a bound on $|\mu(h(W)) - \Phi_\Sigma(h)|$ for three times differentiable h. If this bound is small for all h with some fixed bound on $\|D^k h\|_\infty$, $k = 1, 2, 3$, then W is almost normally distributed.

As seen in (7) the bound on $|\mu(h(W)) - \Phi_\Sigma(h)|$ depends on N and K and the functions ρ and $\tilde{\rho}$. Note that, when K decreases, then the first term in (7) decreases, but the second term increases. Therefore, with fixed N, finding a right balance between small and large values of K yields the smallest value for the bound (7). This bound can then be used to give convergence rate for CLT by choosing $0 \le K < N$ as a function of N. For example, when ρ and $\tilde{\rho}$ decay exponentially, then Corollary 2.2 in **(A)** tells that the choice $K = C \log N$ yields the convergence rate $(\log N)/\sqrt{N}$.

How then to derive the bounds for ρ and $\tilde{\rho}$? The first one is more standard problem, so we focus on the second one. The reason that we might consider that $\tilde{\rho}$ decays at all is that $W^n = \frac{1}{\sqrt{N}}(\sum_{k=0}^{n-K-1} f^k + \sum_{k=n+K+1}^{N-1} f^k)$, where individual terms f^k are only weakly dependent on f^n, if the system in question mixes sufficiently fast. The assumption $\mu(f) = 0$ is also required. An abstract scheme for showing that $\tilde{\rho}(K)$ decays for some concrete model is given in Section 7 of **(A)** and this scheme is then applied to two examples: an angle-doubling map system (with scalar-valued observable) and dispersing billiards model.

In Theorem 2.4 of **(A)** we also derive CLT for continuous time. This theorem uses the previous discrete time theorem and it is the only continuous time result in the articles of this thesis. Thus we omit stating it here to avoid the introduction of new notations.

The previous theorem can be applied for univariate f, but in **(A)** we also proved the following theorem, which gives Wasserstein distance for similar kind of conditions on f.

Theorem 4.2. *Let $f : X \to \mathbb{R}$ be a bounded measurable function with $\mu(f) = 0$. Fix integers $N > 0$ and $0 \le K < N$. Suppose that the following conditions are satisfied:*

(B1) *There exist constants $C_2 > 0$ and $C_4 > 0$, and a non-increasing function $\rho : \mathbb{N}_0 \to \mathbb{R}_+$ with $\rho(0) = 1$ and $\sum_{i=1}^\infty i\rho(i) < \infty$, such that*

$$
\begin{aligned}
|\mu(f\, f^k)| &\le C_2\, \rho(k) \\
|\mu(f\, f^l f^m f^n)| &\le C_4 \min\{\rho(l), \rho(n-m)\} \\
|\mu(f\, f^l f^m f^n) - \mu(f\, f^l)\mu(f^m f^n)| &\le C_4\, \rho(m-l)
\end{aligned}
$$

hold whenever $k \ge 0$ and $0 \le l \le m \le n < N$.

(B2) *There exists a function $\tilde{\rho} : \mathbb{N}_0 \to \mathbb{R}_+$ such that, given a differentiable $A : \mathbb{R} \to \mathbb{R}$ with A' absolutely continuous and $\max_{0 \le k \le 2} \|A^{(k)}\|_\infty \le 1$,*

$$|\mu(f^n A(W^n))| \le \tilde{\rho}(K)$$

holds for all $0 \le n < N$.

(B3) *f is not a coboundary.*

Then

$$\sigma^2 = \mu(f\, f) + 2 \sum_{n=1}^\infty \mu(f\, f^n) \tag{9}$$

is strictly positive and finite. Moreover, if $Z \sim \mathcal{N}(0, \sigma^2)$ is a random variable with normal distribution of mean zero and variance σ^2, then

$$d_{\mathscr{W}}(W, Z) \leq C_{\#}\left(\frac{K+1}{\sqrt{N}} + \sum_{i=K+1}^{\infty} \rho(i)\right) + C'_{\#}\sqrt{N}\tilde{\rho}(K), \tag{10}$$

where

$$C_{\#} = 11 \max\{\sigma^{-1}, \sigma^{-2}\} \max\{C_2, \sqrt{C_4}\}(1 + \|f\|_\infty) \sum_{i=0}^{\infty}(i+1)\rho(i)$$

and

$$C'_{\#} = 2 \max\{1, \sigma^{-2}\}$$

are independent of N and K.

The theorem above is proven with similar arguments as Theorem 4.1. Differences in the conditions and in the statement arise from a different Stein equation, which is $\sigma^2 A'(w) - wA(w) = h(w) - \Phi_{\sigma^2}(h)$ for univariate normal distribution; and from different metric used to compare distribution, while in multivariate case we were bounding $|\mu(h(W)) - \Phi_\Sigma(h)|$ for three times differentiable functions h, we use 1-Lipschitz functions in the place of h in the theorem above. This choice of metric, namely Wasserstein distance, leads to the bound (10) which depends on the variance σ^2, unlike the bound in Theorem 4.1 with no explicit dependence on the covariance matrix Σ.

Note that from (10) we can derive a bound on the commonly used Kolmogorov distance $d_{\mathscr{K}}$ by applying the fact that

$$d_{\mathscr{K}}(W, Z) \leq (2\pi^{-1})^{\frac{1}{4}} \sigma^{-\frac{1}{2}} d_{\mathscr{W}}(W, Z)^{\frac{1}{2}}.$$

However this bound can be far from optimal.

Due to the modifications on measure of distance and underlying Stein equation the distance also depends on the variance σ^2.

Below we give Pène's theorem [60, Theorem 1.1], since it has some resemblance to Theorem 4.1 of this introduction. We will discuss those similarities and differences afterwards. The symbol S_n in the theorem is defined as $\sum_{k=1}^{n} f^k$.

Theorem 4.3. *Let $(f^k)_{k\geq 0}$ be a sequence of stationary R^d-valued bounded random variables defined on $(\Omega, \mathcal{F}, \nu)$ with expectation 0. Let us suppose that there exist two real numbers $C \geq 1, M \geq \max(1, \|f_0\|_\infty)$ and an integer $r \geq 0$ and a sequence of real numbers $(\varphi_{p,l})_{p,l}$ bounded by 1 with $\sum_{p\geq 1} p \max_{l=0,\dots,\lfloor p/(r+1)\rfloor} \varphi_{p,l} < +\infty$ such that for any integers $a, b, c \geq 0$ satisfying $1 \leq a+b+c \leq 3$, for any integers i, j, k, p, q, l with $1 \leq i \leq j \leq k \leq k+p \leq k+p+q \leq k+p+l$, for any $i_1, i_2, i_3 \in \{1, \dots, d\}$, for any $F\colon \mathbb{R}^d \times ([-M; M]^d)^3 \to \mathbb{R}$ bounded, differentiable, with bounded differential, we have*

$$\begin{aligned}&|\operatorname{Cov}(F(S_{i-1}, f^i, f^j, f^k), (f_{(i_1)}^{k+p})^a (f_{(i_2)}^{k+p+q})^b (f_{(i_3)}^{k+p+l})^c)| \\ &\leq C(\|F\|_{L^\infty} + \||DF|_\infty\|_{L^\infty})\varphi_{p,l},\end{aligned}$$

where DF is the Jacobian matrix of F and $f_m^{(s)}$ is the sth coordinate of f_m. Then, the following limit exists:

$$\Sigma := \lim_{n\to+\infty} \frac{1}{n}(\mathbb{E}[S_n^{\otimes 2}]).$$

If $\Sigma = 0$, then the sequence $(S_n)_n$ is bounded in L^2.

Otherwise the sequence of random variables $(\frac{S_n}{\sqrt{n}})_{n\geq 1}$ converges in distribution to a Gaussian random variable N with expectation 0 and with covariance matrix Σ and there *exists a real number $B > 0$ such that, for any integer $n \geq 1$ and any Lipschitz continuous function $\phi\colon \mathbb{R}^d \to \mathbb{R}$, we have*

$$\left|\mathbb{E}\left[\phi\left(\frac{S_n}{\sqrt{n}}\right)\right] - \mathbb{E}[\phi(N)]\right| \leq \frac{BL_\phi}{\sqrt{n}}.$$

Theorem 4.1 of this introduction and Pène's theorem have some similarities and some differences. First of all, both theorems give CLT if some correlation bounds are satisfied. The type of correlation bounds as in (A1) must be satisfied to apply Pène's theorem; for example, choose $F(S_{i-1}, f^i, f^j, f^k) = f^k_\alpha$ and $(f^{k+p}_{(i_1)})^a(f^{k+p+q}_{(i_2)})^b(f^{k+p+l}_{(i_3)})^c = f^{k+p}_\beta$, then

$$\begin{aligned}\nu(f^0_\alpha f^p_\beta) &= \nu(f^0_\alpha f^p_\beta) - \nu(f^0_\alpha)\nu(f^p_\beta) \leq |\nu(f^k_\alpha f^{k+p}_\beta) - \nu(f^k_\alpha)\nu(f^{k+p}_\beta)| \\ &= |\operatorname{Cov}(f^k_\alpha, f^{k+p}_\beta)| \leq C(\|f^k\|_\infty + 1)\varphi_{p,0}.\end{aligned}$$

Furthermore to satisfy conditions on Pène's theorem, a large class of other correlation bounds that are not required on Theorem 4.1 need to be established. On the other hand the type of correlation bound on Condition (A2) is required only in Theorem 4.1. Further differences on these two theorems are that the class of functions that ϕ belongs is larger than the corresponding class of h. A clear benefit of Pène's theorem is that it gives the rate of convergence $O(N^{-1/2})$ instead of $O(N^{-1/2}\log N)$, which results from Theorem 4.1 with exponentially decreasing correlation bounds. However, Stein's method works also on time-dependent setting, which we will take a look next.

5. Time-dependent dynamical system

5.1. **Time dependent systems.** Recall that a time dependent dynamical system (TDDS) is a system with a state space X in which the time evolution is defined by a sequence of transformations $T_1, T_2, \ldots$, where $T_i\colon X \to X$ for all $i \in \mathbb{N}$. Thus if the state of the system at time 0 is x, then at time $t \in \mathbb{N}_0$ it is $T_t \circ T_{t-1} \circ \cdots \circ T_1(x)$. TDDSs are harder to study than SDS, since for example in the general case there is no non-trivial measure μ in X such that it is time invariant with respect to all T_i, $i \in \mathbb{N}$, simultaneously. In this section we assume that these maps T_i are non-random whereas in Section 7 we consider random dynamical systems, in which case the maps T_i are chosen randomly from some probability distribution.

Results in this section are from articles **(B)** and **(C)**. Two general theorems for normal approximations on TDDS setting and a few theorems on an expanding circle map model proven in **(B)** are discussed in this section. We also consider results for Pomeau-Manneville maps proven in article **(C)**.

5.2. **Previous research.** In [44] Lasota and Yorke show that a class of piecewise continuous, piecewise C^1 transformations on the interval $[0, 1]$ have absolutely continuous invariant measures. Conze and Raugi prove TDDS limit theorems in [17]. In [57] Ott, Stenlund and Young proved memory loss properties for time-dependent dynamical systems. We will discuss the result later in Subsection 5.4. In [69] memory-loss properties for Anosov diffeomorphism are studied by Stenlund. Nándori, Szász and Varjú prove a CLT [54, Theorem 1] and apply it to two examples of time-dependent systems on $\mathbb{S}^1$. Haydn et al. [36] prove almost sure invariance principle (ASIP) for several different types of TDDS including β-transformations, perturbed expanding maps of the circle and covering maps.

5.3. **General theorems.** In article **(B)** we prove a similar type of theorem as Theorem 4.1 in this introduction, but this time in the TDDS setting. It turns out that time-dependence does not add serious new difficulties for the proof of the theorem below. Main differences in the proof originate from the lack of invariant measure. First of all, we need to define centered observables $\bar{f}^i = f^i - \mu(f^i)$, where $f^i = f \circ T_i \circ ... \circ T_1$. In the theorem below, sum $W = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \bar{f}^k$ of centered and normalized observables is considered and we define $W_n = W - \frac{1}{\sqrt{N}} \sum_{k \in [n]_K} \bar{f}^k$ analogously to the SDS theorem. In the TDDS setting we compare the distribution of W to $\mathcal{N}(0, \Sigma_N)$, where Σ_N is the covariance matrix of $W(N)$. There might not exist any limit matrix $\lim_{n\to\infty} \Sigma_n$ in the time-dependent setting. Due to lack of invariant measure we can not deduce for example that $\mu(\bar{f}_i \bar{f}_j) = \mu(\bar{f} \bar{f}_{j-i})$, which causes some changes to the formulation of Condition (A1). For more thorough discussion on the differences of the SDS and TDDS version of the theorem, see Section 7 in **(B)**.

Theorem 5.1. *Let $(X, \mathcal{B}, \mu)$ be a probability space and $(f^i)_{i=0}^\infty$ a sequence of random vectors with common upper bound $\|f\|_\infty \geq \|f^i\|_\infty$, for every $i \in \mathbb{N}_0$. Let $h : \mathbb{R}^d \to \mathbb{R}$ be three times differentiable with $\|D^k h\|_\infty < \infty$ for $1 \leq k \leq 3$. Fix integers $N > 0$ and $0 \leq K < N$. Suppose that the following conditions are satisfied:*

(A1) *There exist constants $C_2 > 0$ and $C_4 > 0$, and a non-increasing function $\rho : \mathbb{N}_0 \to \mathbb{R}_+$ with $\rho(0) = 1$ and $\sum_{i=1}^\infty i\rho(i) < \infty$, such that for all $0 \leq i \leq j \leq k \leq l \leq N-1$,*

$$|\mu(\bar{f}^i_\alpha \bar{f}^j_\beta)| \leq C_2 \rho(j-i),$$

$$|\mu(\bar{f}^i_\alpha \bar{f}^j_\beta \bar{f}^k_\gamma \bar{f}^l_\delta)| \leq C_4 \rho(\max\{j-i, l-k\}),$$
$$|\mu(\bar{f}^i_\alpha \bar{f}^j_\beta \bar{f}^k_\gamma \bar{f}^l_\delta) - \mu(\bar{f}^i_\alpha \bar{f}^j_\beta)\mu(\bar{f}^k_\gamma \bar{f}^l_\delta)| \leq C_4 \rho(k-j)$$

hold whenever $k \geq 0$; $0 \leq i \leq j \leq k \leq n < N$; $\alpha, \beta, \gamma, \delta \in \{\alpha', \beta'\}$ and $\alpha', \beta' \in \{1, \dots, d\}$.

(A2) *There exists a function $\tilde{\rho} : \mathbb{N}_0 \to \mathbb{R}_+$ such that*

$$|\mu(\bar{f}^n \cdot \nabla h(v + W_n t))| \leq \tilde{\rho}(K)$$

holds for all $0 \leq n \leq N-1$, $0 \leq t \leq 1$ and $v \in \mathbb{R}^d$.

(A3) *Σ_N is positive-definite $d \times d$ matrix.*

Then

$$|\mu(h(W)) - \Phi_{\Sigma_N}(h)| \leq C_* \left(\frac{K+1}{\sqrt{N}} + \sum_{i=K+1}^\infty \rho(i) \right) + \sqrt{N}\tilde{\rho}(K), \tag{11}$$

where

$$C_* = 6d^3 \max\{C_2, \sqrt{C_4}\} \left(\|f\|_\infty \|D^3 h\|_\infty + \|D^2 h\|_\infty \right) \sqrt{\sum_{i=0}^\infty (i+1)\rho(i)} \tag{12}$$

is independent of N and K.

As in the case of the SDS setting, there also exists a theorem for the Wasserstein distance and a univariate observable f. This theorem is also proven in **(B)** by Stein's method. The differences of this theorem compared to the one above are similar to the differences of Theorems 4.1 and 4.2 of this introduction.

Theorem 5.2. *Let $(X, \mathcal{B}, \mu)$ be a probability space and $(f^i)_{i=0}^{\infty}$ a sequence of random vectors with common upper bound $\|f\|_\infty$. Fix integers $N > 0$ and $0 \le K < N$. Suppose that the following conditions are satisfied.*

(B1) *There exist constants C_2, C_4 and a non-increasing function $\rho : \mathbb{N}_0 \to \mathbb{R}$ with $\rho(0) = 1$, such that for all $0 \le i \le j \le k \le l \le N-1$,*

$$|\mu(\bar{f}^i \bar{f}^j)| \le C_2 \rho(j-i),$$

$$|\mu(\bar{f}^i \bar{f}^j \bar{f}^k \bar{f}^l)| \le C_4 \rho(\max\{j-i, l-k\}),$$
$$|\mu(\bar{f}^i \bar{f}^j \bar{f}^k \bar{f}^l) - \mu(\bar{f}^i \bar{f}^j)\mu(\bar{f}^k \bar{f}^l)| \le C_4 \rho(k-j).$$

(B2) *There exists a function $\tilde{\rho} : \mathbb{N}_0 \to \mathbb{R}_+$ such that, given a differentiable $A : \mathbb{R} \to \mathbb{R}$ with A' absolutely continuous and $\max_{0 \le k \le 2} \|A^{(k)}\|_\infty \le 1$,*

$$|\mu(\bar{f}^n A(W_n))| \le \tilde{\rho}(K)$$

holds for all $0 \le n < N$.

(B3) $\sigma_N^2 > 0$.

Then the Wasserstein distance $d_{\mathscr{W}}(W, \sigma_N Z)$ is bounded from above by

$$C_\# \left(\frac{K+1}{\sqrt{N}} + \sum_{i=K+1}^{\infty} \rho(i) \right) + C'_\# \sqrt{N} \tilde{\rho}(K),$$

where

$$C_\# = 12 \max\{\sigma_N^{-1}, \sigma_N^{-2}\} \max\{C_2, \sqrt{C_4}\}(1 + \|f\|_\infty) \sqrt{\sum_{i=0}^{\infty} (i+1)\rho(i)}$$

and

$$C'_\# = 2 \max\{1, \sigma_N^{-2}\}$$

are independent of K.

5.4. **Expanding circle maps.** In [40] Kawan studies TDDS on compact and connected Riemann manifolds, where the transformations are in C^2 and satisfy a fixed lower bound on the first derivative and an upper bound for the second derivative. He proves that the metric entropy of this system is independent of the initial measure.

With similar assumptions as in [17] Haydn et al. prove an almost sure invariance principle in [36, Theorem 3.1]. Aimino and Rousseau also consider similar setup in [4].

For other papers dealing with expanding circle maps on time-dependent setting; see [53] and [27].

5.4.1. *The model.* We will next introduce the specific model that we examine in **(B)**:

Let the circle $\mathbb{S}^1$ be the state space. Fix $\lambda > 1$ and $A_* > 0$. Denote the set of C^2 expanding circle maps $T : \mathbb{S}^1 \to \mathbb{S}^1$ with the bounds $\inf T' \ge \lambda, \|T''\|_\infty \le A_*$ by $\mathcal{M}$. For the rest of the subsection we assume that all transformations are drawn form the set $\mathcal{M}$. Let $f : \mathbb{S}^1 \to \mathbb{R}^d$ be an Lipschitz continuous observable, i.e., all the coordinate functions of f are Lipschitz continuous. Write $\mathrm{Lip}(f) = \max\{\mathrm{Lip}(f_\alpha) : \alpha \in 1, ..., d\}$ and $\|f\|_{\mathrm{Lip}} = \|f\|_\infty + \mathrm{Lip}(f)$. We also assume that the initial probability measure μ on $\mathbb{S}^1$ has a density ϱ with respect to Lebesgue measure m on $\mathbb{S}^1$ such that $\log \varrho$ is Lipschitz continuous with constant $L_0 = \mathrm{Lip}(\log \varrho)$. We define W, Σ_N and σ_N^2 as before.

In [57] Ott, Stenlund and Young proved memory loss properties for time-dependent dynamical systems in the model $\mathcal{M}$. Theorem 1 in [57] implies that every pair of Lipschitz continuous probability densities φ, ψ converge towards each other exponentially fast. To be more precise, there exists $\nu \in (0,1)$ such that $\int |\mathcal{L}_n(\varphi) - \mathcal{L}_n(\psi)| \, \mathrm{d}m \le C_{\varphi,\psi}\nu^n$ for every $n \in \mathbb{N}_0$, where $\mathcal{L}_n$ is a transfer operator associated to $T_n \circ ... \circ T_1$. These memory loss properties are crucial in bounding the terms in conditions (A2) and (B2) for the model $\mathcal{M}$.

The following theorem from article **(A)** is a CLT with a rate of convergence for TDDS of expanding circle maps and multivariate observable. The distance of two distributions is measured by three times differentiable test functions h. The constants in the following results in this subsection satisfy $\vartheta \in (0,1)$, $C_2, C_4, B_0 > 0$ and depend only on λ, A_*, ϱ and $\|f\|_{\mathrm{Lip}}$.

Theorem 5.3. *Let $(T_i)_{i=1}^{\infty} \subset \mathcal{M}$ be a sequence of transformations in the model $\mathcal{M}$. Let $h\colon \mathbb{R}^d \to \mathbb{R}$ be three times differentiable with $\|D^k h\|_\infty < \infty$, $k = 1,2,3$. Suppose that $N \ge 16/(1-\vartheta)^2$ is such that the matrix Σ_N is positive definite. Then*

$$|\mu(h(W)) - \Phi_{\Sigma_N}(h)| \le C N^{-\frac{1}{2}} \log N,$$

where

$$\begin{aligned} C = & \frac{30 d^3 \max\{C_2, \sqrt{C_4}\} \left(\|f\|_\infty \|D^3 h\|_\infty + \|D^2 h\|_\infty\right)}{(1-\vartheta)^2} \\ & + 2d^2 \|D^2 h\|_\infty \frac{\|f\|_{\mathrm{Lip}}^2}{\vartheta^{-\frac{1}{2}} - \vartheta^{\frac{1}{2}}} + 4dB_0 \|f\|_{\mathrm{Lip}} \|Dh\|_\infty + \frac{2d\|Dh\|_\infty \|f\|_{\mathrm{Lip}}}{\vartheta^{\frac{1}{2}}}. \end{aligned}$$

For univariate f in the same expanding circle TDDS model, we have the following CLT with a rate of convergence in Wasserstein distance:

Theorem 5.4. *Let $(T_i)_{i=1}^{\infty} \subset \mathcal{M}$ be a sequence of transformations in the model $\mathcal{M}$. Let $N \ge 16/(1-\vartheta)^2$ and $\sigma_N \ge C_0 N^{-p}$, where $C_0 > 0$, $p \ge 0$. Then*

$$d_{\mathscr{W}}(W, \sigma_N Z) \le \tilde{C} \max\{1, C_0^{-2}\} N^{-\frac{1}{2}+2p} \log N,$$

where

$$\tilde{C} = \frac{60 \max\{C_2, \sqrt{C_4}\}(1 + \|f\|_\infty)}{(1-\vartheta)^2} + \frac{4\|f\|_{\mathrm{Lip}}^2}{\vartheta^{-\frac{1}{2}} - \vartheta^{\frac{1}{2}}} + 8B_0\|f\|_{\mathrm{Lip}} + \frac{4\|f\|_{\mathrm{Lip}}}{\vartheta^{\frac{1}{2}}} \tag{13}$$

is independent of N.

In particular, if $\sigma_N > C_0$ (case $p = 0$) for $N \ge 3$, the upper bound becomes

$$\tilde{C} \max\{1, C_0^{-2}\} N^{-\frac{1}{2}} \log N.$$

If the variance σ_N is small for large values of N, then the previous theorem is not useful. However, in that case the laws of W and $\sigma_N Z$ are close to Dirac delta distribution in the sense of Wasserstein distance and then estimate $d_{\mathscr{W}}(W, \sigma_N Z) \le 2\sigma_N$ can be used. Since this estimate is strong when σ_N is small, we are able to provide the following CLT result which is independent of variance.

Corollary 5.5. *Let $(T_i)_{i=1}^{\infty}$, $T_i : \mathbb{S}^1 \to \mathbb{S}^1, i \in \mathbb{N}$, be a sequence of transformations in the model $\mathcal{M}$. Then*

$$d_{\mathscr{W}}(W, \sigma_N Z) \leq \max\{\tilde{C}, 2\} N^{-\frac{1}{6}} \log N,$$

for all $N \geq 16/(1-\vartheta)^2$, where $\tilde{C}$ is as in (13).

In the results above the variance σ_N^2 could change as a function of N. We can also study self-normalized version of W, which has variance 1 for all N. For this purpose we define

$$S_N = \sum_{i=0}^{N-1} \bar{f}^i = \sqrt{N} W(N) = \sqrt{N} W \tag{14}$$

and

$$s_N^2 = \mathrm{Var}(S_N) = \mathrm{Var}(\sqrt{N} W) = N \sigma_N^2,$$

Now $S_N/s_N = W/\sigma_N$ has a variance 1 if $s_N > 0$ and we have the following corollary for self-normalized version of W:

Corollary 5.6. *Let $(T_i)_{i=1}^{\infty} \subset \mathcal{M}$ be a sequence of transformations in the model $\mathcal{M}$. Let $N \geq 16/(1-\vartheta)^2$ and $s_N^2 \geq C_0 N^p$, where $C_0 > 0$ and $0 \leq p \leq 1$. Then*

$$d_{\mathscr{W}}\left(\frac{S_N}{s_N}, Z\right) = \tilde{C} \max\{C_0^{-\frac{1}{2}}, C_0^{-\frac{3}{2}}\} N^{1-\frac{3p}{2}} \log N.$$

The corollary above tells us that if the growth of s_N^2 is linear ($p = 1$), then the Wasserstein distance of self-normalized W and $Z \sim \mathcal{N}(0,1)$ is bounded by $CN^{-1/2} \log N$. Furthermore, for $p > 2/3$, we have $d_{\mathscr{W}}(S_N/s_N, Z) \to 0$, when $N \to \infty$.

5.5. **Pomeau-Manneville maps.** The model $\mathcal{M}$ of expanding circle maps satisfies exponential decorrelation, i.e., the functions ρ and $\tilde{\rho}$ in theorems 5.1 and 5.2 of this introduction decay exponentially. This guarantees good convergence rates in the results for expanding circle maps in the previous subsection. However, some dynamical systems do not have such fast decay of correlation. Pomeau-Manneville (P-M) maps are the simplest example of DS with only polynomial decay rates of correlation. Thus these maps form a good starting place to understand DSs with polynomial correlation decay. P-M maps are transformations on the circle, which have a neutral fixed point at $0 \in \mathbb{S}^1$, but are expanding everywhere else. There exists some differences on the definition of P-M maps in the literature; see for example [49], [61] and [2]. We give the exact definition of P-M maps used in **(C)** after first reviewing some literature:

Polynomial loss of memory is proved on [2] for a TDDS with (modified) Pomeau-Manneville maps. Leppänen [47] shows a functional correlation bound for TDDS of intermittent maps, which we also apply in **(C)**. He then proves two multivariate central limit theorems with a rate of convergence for SDS. Combining the results of [47] and **(B)**, we show in **(C)** how multivariate CLT can also be achieved for TDDS with sequence of Pomeau-Manneville maps. For other papers related to topics of this subsection; see [55], [43] and [26].

Following [49], we define for each $\alpha \in (0,1)$ the map $T_\alpha : [0,1] \to [0,1]$ by

$$T_\alpha(x) = \begin{cases} x(1 + 2^\alpha x^\alpha) & \forall x \in [0, 1/2), \\ 2x - 1 & \forall x \in [1/2, 1]. \end{cases} \tag{15}$$

We denote $\mathcal{P} = \{T_\alpha : \alpha \in (0,1)\}$ and call $\mathcal{P}$ a class of Pomeau-Manneville maps.

Each map in $\mathcal{P}$ has a neutral fixed point at the origin. For every $x \in [0,1]$ and $\alpha \in (0,1)$ the derivative of T_α satisfies $T'_\alpha(x) \geq 1$. For small positive values of x, we have $T'_\alpha(x) \approx 1$. Thus T_α expands slowly around the origin. Furthermore for large values of α, the area of slow expansion is larger, while for α close to 0, T_α starts to resemble the angle-doubling map.

Define a convex cone of functions

$$\begin{aligned}\mathcal{C}_*(\alpha) = \{g \in C((0,1]) \cap L^1\ : g \geq 0,\ g \text{ decreasing},\\ x^{\alpha+1} g \text{ increasing},\ g(x) \leq 2^\alpha (2+\alpha) x^{-\alpha} m(g)\},\end{aligned}$$

where $m(g) = \int_0^1 g(x)\,\mathrm{d}x$. Let $\beta_* \in (0,1)$. We call a sequence $(T_{\alpha_n})_{n\geq 1}$ of intermittent maps admissible, if $0 \leq \alpha_n \leq \beta_*$ for all $n \geq 1$.

Let $f\colon [0,1] \to \mathbb{R}^d$. Write

$$W = W(N) = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} (f \circ T_{\alpha_n} \circ \cdots \circ T_{\alpha_1} - \mu(f \circ T_{\alpha_n} \circ \cdots \circ T_{\alpha_1})).$$

The following four results are CLTs with a rate of convergence for time-dependent Pomeau-Manneville systems. The first two theorems are proved via Stein method using general Theorems 5.2 and 5.3. The last two results use similar ideas as Corollaries 5.5 and 5.6.

Theorem 5.7. *Let $(T_{\alpha_n})_{n\geq 1}$ be an admissible sequence of maps. Suppose that the density of the initial measure μ belongs to $\mathcal{C}_*(\beta_*)$, where $\beta_* < 1/3$. Let $N \geq 2$ and let $f : [0,1] \to \mathbb{R}^d$ be a Lipschitz continuous function such that Σ_N is positive definite. Then, for any three times differentiable function $h : \mathbb{R}^d \to \mathbb{R}$ with $\max_{k=1,2,3} \|D^k h\|_\infty < \infty$,*

$$|\mu(h(W)) - \Phi_{\Sigma_N}(h)| \leq C N^{\beta_* - \frac{1}{2}} (\log N)^{\frac{1}{\beta_*}},$$

where $C > 0$ is a constant independent of N. Here $\Phi_{\Sigma_N}(h)$ denotes the expectation of h with respect to $\mathcal{N}(0, \Sigma_N)$.

Theorem 5.8. *Let $(T_{\alpha_n})_{n\geq 1}$ be an admissible sequence of maps. Let $Z \sim \mathcal{N}(0,1)$ be a random variable with normal distribution of mean 0 and variance 1. Suppose that the density of the initial measure μ belongs to $\mathcal{C}_*(\beta_*)$, where $\beta_* < 1/3$. Moreover, let $N \geq 2$ and let $f : [0,1] \to \mathbb{R}$ be a Lipschitz continuous function such that $\sigma_N > 0$. Then,*

$$d_{\mathscr{W}}(W, \sigma_N Z) \leq C \max\{1, \sigma_N^{-2}\} N^{\beta_* - \frac{1}{2}} (\log N)^{\frac{1}{\beta_*}},$$

where $C > 0$ is a constant independent of N.

Proposition 5.9. *Let Z, μ, β_* and f be as in Theorem 5.8. For any $N \geq 2$,*

$$d_{\mathscr{W}}(W, \sigma_N Z) \leq C N^{\frac{2\beta_* - 1}{6}} (\log N)^{\frac{1}{\beta_*}}.$$

Denote

$$S_N = \sqrt{N} W = \sum_{n=0}^{N-1} (f \circ T_{\alpha_n} \circ \cdots \circ T_{\alpha_1} - \mu(f \circ T_{\alpha_n} \circ \cdots \circ T_{\alpha_1})).$$

Then

Corollary 5.10. *Let Z, μ, β_* and f be as in Theorem 5.8. Assume that $Var(S_N) \geq C N^\varepsilon$, where $C > 0$ and $0 \leq \varepsilon \leq 1$. Then, for any $N \geq 2$,*

$$d_{\mathscr{W}}\left(\frac{S_N}{\sqrt{Var(S_N)}}, Z\right) \leq C N^{1 - \frac{3}{2}\varepsilon + \beta_*} (\log N)^{\frac{1}{\beta_*}}.$$

In particular, if $\varepsilon > \frac{2}{3}(1+\beta_*)$*, then*

$$\frac{S_N}{\sqrt{Var(S_N)}} \xrightarrow{d} \mathcal{N}(0,1). \tag{16}$$

In the smaller parameter range $\beta_* < 1/9$, it is seen from [55, Theorem 3.1] that (16) holds for a weaker lower bound on the variance, namely $\mathrm{Var}(S_N) \geq CN^{\varepsilon}$ for some $\varepsilon > 1/2(1-2\beta_*)$ suffices.

6. Quasistatic dynamical system

6.1. **Definition and previous studies.** The concept of quasistatic dynamical system (QDS), was introduced by Stenlund and Dobbs in [19] to model situations where the dynamics change very gradually over time due to weak external forces. The paper describes that the motivation for introducing this concept was to understand the behaviour of a "small" system s whose time-evolution is connected to an ambient system, call it S, such that i) s has virtually no effect on S, and ii) the interaction of S on s is such that s changes very slowly.

For example the Sun (S) can be thought to be in this kind of relationship with the Earth's weather system (s). In [19] a model of dispersing billiards on a torus with (infinitesimally) slowly moving scatterers is given as an example of QDS. Formally QDS is defined as follows:

Definition 6.1 (Discrete time QDS)**.** *Let* $(X, \mathscr{F})$ *be a measurable space,* $\mathcal{S}$ *a topological space whose elements are measurable self-maps* $T : X \to X$*, and* $\mathbf{T}$ *a triangular array of the form*

$$\mathbf{T} = \{T_{n,k} \in \mathcal{S} \ : \ 0 \leq k \leq n, \ n \geq 1\}.$$

If there exists a piecewise continuous curve $\gamma : [0,1] \to \mathcal{M}$ *such that*[3]

$$\lim_{n\to\infty} T_{n,\lfloor nt \rfloor} = \gamma_t \tag{17}$$

for all t*, we say that* $(\mathbf{T}, \gamma)$ *is a* quasistatic dynamical system (QDS) *with* state space X *and* system space $\mathcal{S}$*.*

6.2. **Expanding circle maps.** In the paper [19], where QDS is defined, Dobbs and Stenlund also consider a specific model, with a state space $\mathbb{S}^1$ and a system space $\mathcal{M}$ defined in Subsection 5.4 of this introduction. The topology of $\mathcal{M}$ is defined by a metric d_{C^1}, where $d_{C^1}(T_1, T_2) = \sup_{x\in\mathbb{S}^1} d(T_1(x), T_2(x)) + \|T_1' - T_2'\|_\infty$ and d is the natural metric on $\mathbb{S}^1$. The curve γ is assumed to be Hölder continuous with exponent $\eta \in (0,1)$. Furthermore it is assumed that the maps in the array $\mathbf{T}$ satisfy:

$$\sup_{n\geq 1} n^{\eta} \sup_{t\in[0,1]} d_{C^1}(T_{n,\lfloor nt\rfloor}, \gamma_t) < \infty.$$

Denote $f_{n,k} = f \circ T_{n,k} \circ \cdots \circ T_{n,1}$ and $S_n(x,t) = \int_0^{nt} f_{n,\lfloor ns \rfloor}(x) \mathrm{d}s$. By fixing $t \in [0,1]$ we can think of $S_n(x,t)$ as a Birkhoff sum. Denote $\zeta_n(x,t) = n^{-1}S_n(x,t)$. For each $x \in \mathbb{S}^1$ there exist a corresponding map in $C^0([0,1], \mathbb{R})$ defined by $t \mapsto \zeta_n(x,t)$. Given an initial probability measure μ on $\mathbb{S}^1$ this map defines a probability distribution $\mathbf{P}_n^{\mu}$ on $C^0([0,1], \mathbb{R})$.

We denote the invariant SRB measure with respect to γ_t by $\hat{\mu}_t$ and $\hat{\mu}_t(f) = \int f \, \mathrm{d}\hat{\mu}_t$. Now [19, Theorem 3.1] states the following:

[3]For any real number $s \geq 0$, $\lfloor s \rfloor$ denotes the integer part of s.

Theorem 6.2. *Suppose f is Lipschitz continuous and μ is absolutely continuous. The function $t \mapsto \hat{\mu}_t(f)$ is continuous. The measures $\mathbf{P}_n^\mu$ converge weakly, as $n \to \infty$, to the point mass at $\zeta \in C^0([0,1],\mathbb{R})$, where*

$$\zeta(t) = \int_0^t \hat{\mu}_s(f)\,\mathrm{d}s$$

This means that the stochastic process $\zeta_n(t)$ converges to the non-random limit $\zeta(t)$, when $n \to \infty$. Note that the theorem above can be generalized for γ which is only piecewise Hölder continuous with finitely many discontinuities; see Theorem 3.10 in the same paper.

Although Birkhoff ergodic theorem is not applicable in this QDS setting, we can formulate similar type of question, namely: Let $t \in [0,1]$ be fixed. Does $n^{-1}S_n(x,t)$ converge to some limit for almost every x, with respect to some measure, and if it does, how to compute that limit? As the previous theorem suggest the correct limit is $\zeta(t)$. This result, which holds for a piecewise Hölder continuous γ, is given in [70, Theorem 3.3] as follows:

Theorem 6.3. *Let f be continuous. Then*

$$\lim_{n\to\infty} \sup_{t\in[0,1]} |\zeta_n(x,t) - \zeta(t)| = 0, \tag{18}$$

for almost every x in the sense of Lebesgue.

Theorem 2.2 in the same article gives sufficient conditions for (18) to hold for arbitrary QDS. The conditions are very natural and we omit stating the theorem here.

Now when the question concerning the convergence of ζ_n has been solved, we may ask how $n^{1/2}\zeta_n = n^{-1/2}S_n$ behaves. For this purpose we need centering. The question of proper centering is dealt in [19] by defining the concept of admissible centering. In this introduction we are only interested in the fact that $\mu(\zeta_n(\cdot,t))$ is admissible if μ has Lipschitz continuous density. Define $\bar{f}_{n,t} = f_{n,t} - \mu(f_{n,t})$ and write

$$\xi_n(x,t) = n^{1/2}\zeta_n(x,t) - n^{1/2}\mu(\zeta_n(\cdot,t)) = n^{1/2}\int_0^t \bar{f}_{n,\lfloor nr\rfloor}(x)\,\mathrm{d}s$$

Now each ξ_n defines a distribution in $C^0([0,1],\mathbb{R})$ which we denote by $\mathbb{P}_n^\mu$. We also write

$$\hat{f}_t = f - \hat{\mu}_t(f)$$

and the limit of the variance of $(1/\sqrt{m})\sum_{k=0}^{m-1} f\circ\gamma_t^k$ with respect to the measure $\hat{\mu}_t$ by

$$\hat{\sigma}_t^2(f) = \lim_{m\to\infty}\hat{\mu}_t\left[\left(\frac{1}{\sqrt{m}}\sum_{k=0}^{m-1}\hat{f}_t\circ\gamma_t^k\right)^2\right]$$

Now Lemma 3.5.i) and Theorem 3.6 in [19] imply (for the choice $\mu(\zeta_n(\cdot,t))$ of centering):

Theorem 6.4. *Suppose that f and the density of μ are Lipschitz continuous. Then the function $t \to \hat{\sigma}^2(f)$ is continuous and the measures $\mathbb{P}_n^\mu$ converge weakly, as $n \to \infty$, to the law of the process*

$$\xi(t) = \int_0^t \hat{\sigma}_s(f)\,\mathrm{d}W_s.$$

Here W is standard Brownian motion and the stochastic integral is to be understood in the sense of Itō.

The previous theorem yields that $\xi_n(t)$ converges to the Gaussian process ξ. It can also be generalized for piecewise Hölder continuous γ with finitely many discontinuities. A new result in paper **(B)** gives a convergence rate result for fixed t in the same setting, where it is assumed that γ is a Hölder continuous curve (with no discontinuities). In the theorem we denote $\sigma_t^2 = \int_0^t \hat{\sigma}_s^2(f)\,\mathrm{d}s$.

Theorem 6.5. *Let $t_0 \in]0,1]$ be such that $\hat{\sigma}_{t_0}^2 > 0$. Then for all $\eta' < \eta$ there exist constants C such that for every $t \geq t_0$ and $n \geq 1$*

$$d_{\mathscr{W}}(\xi_n(t), \sigma_t Z) \leq Cn^{-\eta'} + Cn^{-\frac{1}{2}} \log n.$$

The following result, also proven in **(B)** assuming that γ is Hölder continuous with no discontinuities, yields a convergence rate without assuming restrictions on variance $\hat{\sigma}_t^2$.

Theorem 6.6. *Let $\eta' < \eta$. Then there exists a constant C such that the following holds for every $t \in [0,1]$ and $n \geq 1$:*

$$d_{\mathscr{W}}(\xi_n(t), \sigma_t Z) \leq Cn^{-\frac{\eta'}{2}} + Cn^{-\frac{1}{6}} \log n.$$

A result for multivariate variables is also proven in **(B)**. To this end we introduce two notations. The covariance matrix of $\xi_n(t)$ with respect to μ is denoted by

$$\Sigma_{n,t} = \mu[\xi_n(t) \otimes \xi_n(t)], \quad n \geq 1,\, t \in [0,1],$$

and we write

$$\hat{\Sigma}_t(f) = \lim_{m\to\infty} \hat{\mu}_t \left[\frac{1}{\sqrt{m}} \sum_{k=0}^{m-1} \hat{f}_t \circ \gamma_t^k \otimes \frac{1}{\sqrt{m}} \sum_{k=0}^{m-1} \hat{f}_t \circ \gamma_t^k \right],$$

given that the limit exists. Furthermore we denote $\Sigma_t = \int_0^t \hat{\Sigma}_s^2(f)\,\mathrm{d}s$.

Note that an analogous result to Theorem 6.4 for multivariate variables is given in [19, Theorem 3.9]. The following result in **(B)** is a multivariate version of Theorem 6.5.

Theorem 6.7. *Let $t_0 \in]0,1]$ be such that $\hat{\Sigma}_{t_0}$ is positive definite and let $h : \mathbb{R}^d \to \mathbb{R}$ be three times differentiable with $\|D^k h\|_\infty < \infty$ for $1 \leq k \leq 3$. Then for all $\eta' < \eta$ there exists constant C independent of t such that for every $t \geq t_0$ and $n \geq 1$*

$$|\mu(h(\xi_n(t))) - \Phi_{\Sigma_t}(h)| \leq Cn^{-\eta'} + Cn^{-\frac{1}{2}} \log n.$$

6.3. **Pomeau-Manneville maps.** Recall the definition of QDS $(\mathbf{T}, \gamma)$ in Subsection 6.1. The intermittent version of the QDS was introduced in [46]. It is a QDS which satisfies the following definition.

Definition 6.8 (Intermittent QDS)**.** *Let $X = [0,1]$ and $\mathcal{P}$[4] the class of Pomeau-Manneville maps (equipped, say, with the uniform topology). Next, let*

$$\{\alpha_{n,k} \in [0,1) \;:\; 0 \leq k \leq n,\; n \geq 1\}$$

be a triangular array of parameters and

$$\tau : [0,1] \to [0,1]$$

a piecewise continuous curve satisfying

$$\lim_{n\to\infty} \alpha_{n,\lfloor nt \rfloor} = \tau_t$$

for all t. Finally, define $\gamma_t = T_{\tau_t}$ and

$$T_{n,k} = T_{\alpha_{n,k}}.$$

[4] recall the definition of $\mathcal{P}$ in Subsection 5.5

In P-M maps setting Leppänen and Stenlund proved the following result [46, Theorem 1.4] that is similar to the expanding circle maps Theorem 6.3 in this introduction. It can be thought as a counterpart of the ergodic theorem for Pomeau-Manneville QDS. In the following theorems $\hat{\mu}_t$ stands for invariant SRB measure associated to T_{τ_t}.

Theorem 6.9. *Suppose that the curve $\tau\colon [0,1] \to [0,1]$ is piecewice Hölder continuous with exponent $\theta \in (0,1]$, that*

$$\overline{\tau([0,1])} \subset [0, \beta_*] \tag{19}$$

for some $\beta_* \in (0,1)$, and that

$$\lim_{n\to\infty} n^{\theta} \sup_{t\in[0,1]} |\alpha_{n,\lfloor nt\rfloor} - \tau_t| < \infty.$$

(i) *If $\beta_* \geq 1/2$, then for each $f \in C([0,1])$,*

$$\lim_{n\to\infty} \sup_{t\in[0,1]} |\zeta_n(x,t) - \zeta(t)| = 0 \tag{20}$$

in probability, with respect to the Lebesgue measure. That is,

$$\lim_{n\to\infty} m\left(\sup_{t\in[0,1]} |\zeta_n(x,t) - \zeta(t)| \geq \epsilon\right) = 0$$

for all $\epsilon > 0$.

(ii) *If $\beta_* < 1/2$, then (20) holds for almost every $x \in [0,1]$ with respect to the Lebesgue measure. The one-parameter family of measures $\mathscr{P} = (\hat{\mu}_t)_{t\in[0,1]}$ is essentially unique in this sense.*

Let μ be the initial probability measure on $(\mathbb{S}^1, \mathcal{B})$, where $\mathcal{B}$ is the Borel σ-algebra on $\mathbb{S}^1$, and let ν be any measure on $(\mathbb{S}^1, \mathcal{B})$. Leppänen [45] defines a fluctuation $\chi_n^{\nu}\colon [0,1]\times[0,1] \to \mathbb{R}$ by

$$\chi_n^{\nu}(x,t) = n^{1/2}\zeta_n(x,t) - n^{1/2}\nu(\zeta_n(\cdot,t)).$$

Now $\chi_n^{\nu}(x,t)$ is defined exactly in the same way as $\xi_n(x,t)$ except that $\chi_n^{\nu}(x,t)$ is centered by ν, while $\xi_n(x,t)$ is centered by the initial measure μ. Given an initial probability measure μ, the map $x \mapsto \chi_n^{\nu}(x,\cdot)$ is a random element with values in $C([0,1])$, and its distribution is denoted by $\mathbb{P}_n^{\mu,\nu}$ in the following theorem from [45], which shows that when certain conditions are satisfied, then $\chi_n^{\nu}(x,t)$ converges to Gaussian process χ for parameter range $\beta_* < 1/2$. The tightness assumption in the theorem is not necessary, since it is implied[5] by [71, Theorem 4.7].

Theorem 6.10. *Let $f\colon [0,1] \to \mathbb{R}$ be Lipschitz continuous, and let the initial measure μ be such that its density belongs to $C_*(\beta_*)$. Suppose that $\tau\colon [0,1] \to [0,1)$ is Hölder-continuous of order $\eta \in (0,1]$, that $\tau([0,1]) \subset [0,\beta_*]$ for some $\beta_* < 1/2$, and that*

$$\lim_{n\to\infty} n^{\eta} \sup_{t\in[0,1]} |\alpha_{n,\lfloor nt\rfloor} - \tau_t| < \infty.$$

Then, the variance $\hat{\sigma}_t^2(f)$ is finite and satisfies the Green-Kubo formula

$$\hat{\sigma}_t^2(f) = \hat{\mu}_t[\hat{f}_t^2] + 2\sum_{k=1}^{\infty} \hat{\mu}_t[\hat{f}_t\hat{f}_t \circ T_{\tau_t}^k].$$

[5] personal communication with Juho Leppänen

If the sequence of measures $(\mathbb{P}_n^{\mu,\mu})_{n\geq 1}$ is tight, then for any probability measure ν, whose density $g = g_1 - g_2$ for some $g_1, g_2 \in C_*(\beta_*)$, the sequence $(\mathbb{P}_n^{\mu,\nu})_{n\geq 1}$ converges weakly to the law of the process

$$\chi(t) = \int_0^t \hat{\sigma}_s(f) dW_s. \tag{21}$$

Here W is a standard Brownian motion, and the stochastic integral is to be understood in the sense of Itō.

In article **(C)** we prove the following multivariate CLT with a rate of convergence for P-M maps in parameter range $\beta_* < 1/3$:

Theorem 6.11. *Let $f : [0,1] \to \mathbb{R}^d$ be a Lipschitz continuous function, and let the initial measure μ be such that its density is in $\mathcal{C}_*(\beta_*)$. Suppose that the limiting curve $\tau : [0,1] \to [0,1]$ is Hölder-continuous of order $\eta \in (0,1]$, that $\tau([0,1]) \subset [0, \beta_*]$ for some $\beta_* < 1/3$, and that*

$$\sup_{n\geq 1} n^\eta \sup_{t\in[0,1]} |\alpha_{n,\lfloor nt \rfloor} - \tau_t| < \infty. \tag{22}$$

Suppose there exists $t_0 \in (0,1]$ such that $\hat{f}_{t_0}$ is not a co-boundary for γ_{t_0} in any direction. Then for all $t \geq t_0$, Σ_t is positive definite, and for all three times differentiable functions $h : \mathbb{R}^d \to \mathbb{R}$ with $\max_{k=1,2,3} \|D^k h\|_\infty < \infty$,

$$|\mu [h(\xi_n(t))] - \Phi_{\Sigma_t}(h)| \leq C n^{-\theta}, \tag{23}$$

where $C > 0$ is independent of t, and

$$\theta = \frac{1}{\frac{12}{\eta(1-\beta_*)} + 1}.$$

This theorem is proven by applying general result also proven in **(C)**. This general result is explained in the next subsection.

6.4. **Main theorem.** In **(C)** we introduce four abstract conditions **(I)** – **(IV)** under which a multivariate CLT for (any, not just intermittent) QDS holds. Conditions **(I)** and **(II)** guarantee the convergence of variance $\lim_{n\to\infty} \Sigma_{n,t} = \Sigma_t$ as shown in Theorem 2.8 of **(C)**. If conditions **(III)** and **(IV)** are satisfied, then a multivariate CLT with a convergence rate holds, which is shown in Theorem 2.9 of **(C)**. We give theorems 2.8 and 2.9 after introducing and discussing conditions **(I)**–**(IV)**:

First set $\mathbf{T}' = \mathbf{T} \cup \{\gamma_t : t \in [0,1]\}$ and $\mathcal{C} = \bigcup_{k=0}^\infty C_k$, where

$$C_0 = \{\hat{\mu}_t, \mu : t \in [0,1]\},$$
$$C_{k+1} = \{(T)_*\nu : \nu \in C_k, T \in \mathbf{T}'\}.$$

Below $\mathcal{T}_k$ stands for any k-composition $T_k \circ \cdots \circ T_1$ of maps $T_i \in \mathbf{T}'$. We assume the existence of a constant $C > 0$, such that the following conditions hold for all bounded functions F of the form $F = f_a \cdot f_b^q \circ T_k \circ \cdots \circ T_1$ where $T_i \in \mathbf{T}'$, $a, b \in \{1, \ldots, d\}$, $q \in \{0, 1\}$:

- **(I)** There is $\varphi > 1$ and $n_1 \in \mathbb{N}$, such that for any $\nu^1, \nu^2 \in \mathcal{C}$, and any $n, m \in \mathbb{N}_0$ with $m - n \geq n_1$,

  $$|\nu^1(f_\alpha^p \circ \mathcal{T}_n \cdot F \circ \mathcal{T}_m) - \nu^1(f_\alpha^p \circ \mathcal{T}_n)\nu^2(F \circ \mathcal{T}_m)| \leq C(m-n)^{-\varphi},$$

  whenever $\alpha \in \{1, \ldots, d\}$ and $p \in \{0, 1\}$.

(II) There is $\psi \in (0,1]$, such that for all $k, m, n \in \mathbb{N}_0$ with $k+m \leq n$, measures $\nu \in \mathcal{C}$, $s, r_1, \ldots, r_k \in [0,1]$, $\alpha \in \{1, \ldots, d\}$, and $p \in \{0,1\}$:

$$\left|\nu\left[f_\alpha^p \cdot (F \circ T_{n,m+k} \circ \cdots \circ T_{n,m+1} - F \circ \gamma_{(m+k)/n} \circ \cdots \circ \gamma_{(m+1)/n})\right]\right| \leq Ckn^{-\psi},$$

and

$$|\hat{\mu}_s[f_\alpha^p \cdot (F \circ \gamma_s^k - F \circ \gamma_{r_k} \circ \cdots \circ \gamma_{r_1})]| \leq Ck \max_{1 \leq l \leq k} |s - r_l|^\psi.$$

(III) There exists $\zeta \in (0,1)$ such that for every $n \in \mathbb{N}$ and $t \geq t_0$,

$$\left|\mu\left[h(\xi_n(\lceil nt \rceil/n))\right] - \Phi_{\Sigma_{n,\lceil nt \rceil/n}}(h)\right| \leq Cn^{-\zeta}.$$

(IV) $\hat{f}_{t_0}$ is not a co-boundary for γ_{t_0} in any direction.[6]

We summarize the conditions here very briefly. Condition **(I)** yields a polynomial memory loss property, when $p = 0$. It also gives a polynomial correlation decay estimate, when we choose $\nu^1 = \nu^2$. Condition **(II)** is a perturbation estimate. It states that, when n is large, changing $T_{n,l}$ to $\gamma_{l/n}$ does not affect much the integral with respect to ν. Similarly the second part of condition **(II)** tells that changing γ_s to γ_r, where $|r - s| \ll 1$ does not make a large difference to the corresponding integral. Overall, the purpose of conditions **(I)** and **(II)** is to guarantee that $|\Sigma_{n,t} - \Sigma_t|$ has a polynomial componentwise upper bound. Condition **(IV)** guarantees that Σ_t is positive definite for $t \geq t_0$. Condition **(III)** then implies that $\xi_n(nt/n)$ is almost normally distributed.

Theorems 2.8 and 2.9 of **(C)** are given below:

Theorem 6.12. *Suppose that conditions (I) and (II) hold. Then, given any $\varepsilon > 0$,*

$$[\Sigma_{n,t}]_{\alpha\beta} - [\Sigma_t]_{\alpha\beta} = \mathcal{O}(n^{\max\{(\psi - \varphi\psi)/(\varphi+\psi+1)+\epsilon, -1/6\}}),$$

for every $\alpha, \beta \in \{1, 2, ..., d\}$.

Theorem 6.13. *Let $h : \mathbb{R}^d \to \mathbb{R}$ be a Lipschitz continuous function and $t_0 \in (0,1]$. Suppose that conditions **(I)**–**(IV)** hold. Then, for every $t \geq t_0$, Σ_t is positive definite and*

$$|\mu\left[h(\xi_n(t))\right] - \Phi_{\Sigma_t}(h)| \leq Cn^{\max\{(\psi - \varphi\psi)/(\varphi+\psi+1)+\epsilon, -1/6, -\zeta\}}$$

holds. Here $\epsilon > 0$ can be chosen arbitrary small, and the constant $C > 0$ depends on t_0 but not on t.

7. Random dynamical system

7.1. **Random dynamical system.** The following definition of random dynamical system (RDS) is from Arnold [6]. The time $\mathbb{T}$ in the definition stands for $\mathbb{R}, \mathbb{R}^+, \mathbb{R}^-, \mathbb{Z}, \mathbb{N}_0, -\mathbb{N}_0$ or $\mathbb{N}$. In our paper **(D)** we only consider the case $\mathbb{T} = \mathbb{N}_0$

Definition 7.1. *A measurable random dynamical system on the measurable space $(X, \mathcal{B})$ over (or covering, or extending) a metric dynamical system $(\Omega, \mathcal{F}, \mathbb{P}, (\theta(t))_{t \in \mathbb{T}})$ with time $\mathbb{T}$ is a mapping*

$$\varphi\colon \mathbb{T} \times \Omega \times X \to X, \qquad (t, \omega, x) \mapsto \varphi(t, \omega, x),$$

with the following properties:

(i) Measurability: φ is $\mathcal{B}(\mathbb{T}) \otimes \mathcal{F} \otimes \mathcal{B}$, $\mathcal{B}$-measurable.

[6]i.e. there is no unit vector $v \in \mathbb{R}^d$ and a function $g_v : X \to \mathbb{R}$ in $L^2(\mu)$ such that $v \cdot f = g_v - g_v \circ \gamma_{t_0}$.

(ii) Cocycle property: The mappings $\varphi(t,\omega) = \varphi(t,\omega,\cdot)\colon X \to X$ form a cocycle over $\theta(\cdot)$, i.e. they satisfy

$$\varphi(0,\omega) = \mathrm{id}_X \quad \textit{for all} \quad \omega \in \Omega, \quad (\textit{if } 0 \in \mathbb{T}),$$
$$\varphi(t+s,\omega) = \varphi(t,\theta(s)\omega) \circ \varphi(s,\omega) \quad \textit{for all} \quad s,t \in \mathbb{T}, \quad \omega \in \Omega.$$

The basic idea of RDS is that the trajectory of the point $x \in X$ is defined by randomly picked self-maps on X. These maps are chosen randomly from a measurable space $(\Omega, \mathcal{F})$ according to the distribution $\mathbb{P}$. If for example $\mathbb{T} = \mathbb{N}_0$ and an element ω is picked from Ω, then the corresponding trajectory of x is the sequence $x, \varphi(1,\omega,x), \varphi(2,\omega,x), \cdots$.

In our articles we study the type of RDS, where $(\Omega, \mathcal{F}, \mathbb{P}) = (\Omega_0^{\mathbb{N}}, \mathcal{E}^{\mathbb{N}}, \mathbb{P})$ in which $(\Omega_0, \mathcal{E})$ is a measurable space. Thus $\omega = (\omega_n)_{n\geq 1}$, where $\omega_t \in \Omega_0$ for every $t \in \mathbb{N}$. We define

$$\tau\colon \Omega \to \Omega : \omega = (\omega_1, \omega_2, \dots) \mapsto \tau\omega = (\omega_2, \omega_3, \dots)$$

and assume that τ is $\mathcal{F}$-measurable, but does not necessarily preserve $\mathbb{P}$. Let $(X, \mathcal{B})$ be a measurable space and $T_{\omega_0}\colon X \to X$ a measurable self-map for every $\omega_0 \in \Omega_0$. We now see that the above definition of RDS is satisfied by choosing $\theta(t) = \tau^t$ and

$$\varphi\colon \mathbb{N}_0 \times \Omega \times X \to X : (t,\omega,x) \mapsto T_{\omega_t} \circ \cdots \circ T_{\omega_1}(x).$$

7.2. **Quenched limit theorems.** By quenched limit theorems we mean the type of limit theorems that give almost sure information about fiberwise centered Birkhoff sums. What this means in the context of our article **(D)**, will be made more precise soon. A lot of work has been done in the research of quenched limit theorems. We point to the following papers [1, 3, 8–11, 14, 20, 21, 27, 41, 54, 55]

In this subsection we examine the variance of normalized and fiberwise centered Birkhoff sums. In article **(D)** we give a set of assumptions under which the limit of this variance is the same for almost every fiber. A formula for computing this variance is also given under these assumptions. We also establish conditions for the variance to be non-zero. These results are applied to yield CLTs in articles **(B)** and **(C)** for expanding circle maps and P-M maps, correspondingly. These applications will be discussed in the last two subsections of this section. First we give some basic notations.

Let $f : X \to \mathbb{R}$ be an observable. Define

$$f_i = f_i(\omega) = f \circ T_{\omega_i} \circ \cdots \circ T_{\omega_1} = f \circ \varphi(i,\omega)$$

and

$$W_n = W_n(\omega) = \frac{1}{\sqrt{n}} \sum_{i=0}^{n-1} f_i.$$

Given an initial probability measure μ on X, we write $\bar{f}_i$ and $\bar{W}_n$ for the corresponding fiberwise-centered random variables:

$$\bar{f}_i(\omega) = f_i - \mu(f_i) \quad \text{and} \quad \bar{W}_n(\omega) = W_n - \mu(W_n).$$

Finally we define the variance

$$\sigma_n^2(\omega) = \mathrm{Var}_\mu \bar{W}_n = \frac{1}{n} \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \mu(\bar{f}_i \bar{f}_j).$$

Our main result in article **(D)** considers estimating $|\sigma_n^2(\omega) - \sigma^2|$, where $\sigma^2 = \lim_{n\to\infty} \sigma_n^2(\omega)$. This limit exists for almost every ω and is non-random, when the soon to be specified Assumptions (SA1)–(SA4) hold. This result is proven by first estimating $|\sigma_n^2(\omega) - \mathbb{E}\sigma_n^2|$ and then $|\mathbb{E}\sigma_n^2 - \sigma^2|$. This main result shows the almost sure convergence of the variance. If it can be shown that $d(\bar{W}_n, \sigma_n Z) \to 0$ where d is for example the Wasserstein distance, then a quenched CLT can be obtained. A quenched CLT is shown for two models: expanding circle maps in article **(B)** and P-M maps in article **(C)**. All these results will be given in this introduction. Beside these, we also show some conditions for the variance σ^2 to be zero or non-zero.

Now we give assumptions (SA1)–(SA4), which are used in the formulation of the main theorem of article **(D)**:

Standing Assumption (SA1). The observable f is a bounded measurable function and μ is a probability measure and a uniform decay of correlations holds in that

$$|\mu(\bar{f}_i \bar{f}_j)| \le \eta(|i-j|)$$

almost surely, where $\eta : \mathbb{N}_0 \to [0, \infty)$ is such that

$$\sum_{i=0}^{\infty} \eta(i) < \infty \quad \text{and} \quad \eta \text{ is non-increasing.} \tag{24}$$

■

Let $\mathcal{F}_1^i \subset \mathcal{F}$ be a sigma-algebra generated by the projections $\pi_1, ..., \pi_i$, where $\pi_k(\omega) = \omega_k$, and $\mathcal{F}_{i+n}^{\infty} \subset \mathcal{F}$ be generated by $\pi_{i+n}, \pi_{i+n+1}, \cdots$.

Define

$$\alpha(\mathcal{F}_1^i, \mathcal{F}_j^{\infty}) = \sup_{A \in \mathcal{F}_1^i,\, B \in \mathcal{F}_j^{\infty}} |\mathbb{P}(AB) - \mathbb{P}(A)\,\mathbb{P}(B)|.$$

In the following $(\alpha(n))_{n \ge 1}$ will denote a sequence such that

$$\sup_{i \ge 1} \alpha(\mathcal{F}_1^i, \mathcal{F}_{i+n}^{\infty}) \le \alpha(n)$$

for each $n \ge 1$.

The next assumption is a strong-mixing condition for (deterministic) dynamical system (Ω, τ). Note that $\alpha \colon \mathbb{N} \to \mathbb{R}$ in the assumption is not the standard strong mixing coefficient function, but any non-increasing upper bound on it.

Standing Assumption (SA2). Random selection process is **strong mixing**: $\alpha(n)$ can be chosen so that

$$\lim_{n \to \infty} \alpha(n) = 0 \quad \text{and} \quad \alpha \text{ is non-increasing.}$$

■

The pushforward of a map T is denoted by T_*, i.e., T_* is the map acting on probability measures m that satisfies $(T_* m)(A) = m(T^{-1}A)$ for measurable sets A. Write

$$\mu_k = (T_{\omega_k} \circ \cdots \circ T_{\omega_1})_* \mu$$

and

$$\mu_{k,r+1} = (T_{\omega_k} \circ \cdots \circ T_{\omega_{r+1}})_* \mu$$

for $k \ge r$. We also write

$$f_{l,k+1} = f \circ T_{\omega_l} \circ \cdots \circ T_{\omega_{k+1}} = f \circ \varphi(l-k, \tau^k \omega)$$

for $l \ge k$. Note that all of these objects depend on ω through the maps T_{ω_i}. We use the conventions $\mu_0 = \mu$, $\mu_{r,r+1} = \mu$ and $f_{k,k+1} = f$ here.

Standing Assumption (SA3). The following uniform **memory-loss condition** holds: there exists a constant $C \geq 0$ such that

$$|\mu_k(g) - \mu_{k,r+1}(g)| \leq C\eta(k-r) \tag{25}$$

for all

$$g \in \mathcal{G}_k = \mathcal{G}_k(\omega) = \{f_{l,k+1} : l \geq k\} \cup \{f f_{l,k+1} : l \geq k\}$$

whenever $k \geq r$. The bound holds uniformly for (almost) all ω. ∎

The idea behind (SA3) is that, while it is satisfied, the transformations $T_{\omega_1}, ..., T_{\omega_{k-r}}$ do not have much effect on the value of $\mu(\bar{f}_k \bar{f}_l)$, when $k - r$ is large.

For the rest of the section we assume that $\mathbb{P}$ is asymptotically mean stationary, with mean $\bar{\mathbb{P}}$. In other words, there exists a measure $\bar{\mathbb{P}}$ such that, given a bounded measurable $g : \Omega \to \mathbb{R}$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=0}^{n-1} \int g \circ \tau^i \, \mathrm{d}\mathbb{P} = \int g \, \mathrm{d}\bar{\mathbb{P}}. \tag{26}$$

The measure $\bar{\mathbb{P}}$ is then τ-invariant. We denote $\bar{\mathbb{E}} g = \int g \, \mathrm{d}\bar{\mathbb{P}}$. We use the following notation in the next assumption:

$$g^1_{ik}(\omega) = \mu(f_i f_{i+k}) = \mu(f \circ \varphi(i,\omega) f \circ \varphi(i+k,\omega))$$

and

$$g^2_{ik}(\omega) = \mu(f_i)\mu(f_{i+k}) = \mu(f \circ \varphi(i,\omega))\mu(f \circ \varphi(i+k,\omega)).$$

Standing Assumption (SA4). The probability measure $\mathbb{P}$ is asymptotically mean stationary, and there exist $C_0 > 0$ and $\zeta > 0$ such that

$$\sup_{r,k,a} \left| \frac{1}{n} \sum_{i=0}^{n-1} \int g^a_{rk} \circ \tau^i \, \mathrm{d}\mathbb{P} - \int g^a_{rk} \, \mathrm{d}\bar{\mathbb{P}} \right| \leq C_0 n^{-\zeta} \tag{27}$$

for all $n \geq 1$. Here the sup is taken over all $r, k \geq 0$ and $a \in \{1, 2\}$. ∎

Note that if $\mathbb{P}$ is stationary then (26) and (27) hold trivially with $\bar{\mathbb{P}} = \mathbb{P}$.

We can now present the main result of article **(D)**. The theorem shows that the variance $\sigma_n^2(\omega)$ converges for almost every ω, when Assumptions (SA1)–(SA4) hold with sufficiently fast decreasing η and α. The formula of the limit $\sigma^2 = \lim_{n\to\infty} \sigma_n^2(\omega)$ is given as is the speed of convergence, which depends on η, α and ζ.

Theorem 7.2. *Assume (SA1&3) with $\eta(n) = Cn^{-\psi}$, $\psi > 1$; (SA2) with $\alpha(n) = Cn^{-\gamma}$, $\gamma > 0$; and (SA4) with $\zeta > 0$. Fix an arbitrarily small $\delta > 0$. Then there exists $\Omega_* \subset \Omega$, $\mathbb{P}(\Omega_*) = 1$, such that all of the following holds: The non-random number*

$$\sigma^2 = \sum_{k=0}^{\infty} (2 - \delta_{k0}) \lim_{i\to\infty} \bar{\mathbb{E}}[\mu(f_i f_{i+k}) - \mu(f_i)\mu(f_{i+k})]$$

is well defined, nonnegative, the series is absolutely convergent, and

$$\lim_{n\to\infty} \sigma_n^2(\omega) = \sigma^2$$

for every $\omega \in \Omega_*$. Moreover, the absolute difference

$$\Delta_n(\omega) = |\sigma_n^2(\omega) - \sigma^2|$$

has the following upper bounds, for any $\omega \in \Omega_*$:

$$\Delta_n = \begin{cases} O(n^{-\frac{1}{2}} \log^{\frac{3}{2}+\delta} n), & \zeta \geq 1, \quad \min\{\psi - 1, \gamma\} > 1, \\ O(n^{-\frac{1}{2}+\delta}), & \zeta \geq 1, \quad \min\{\psi - 1, \gamma\} = 1, \\ O(n^{-\frac{\min\{\psi-1,\gamma\}}{2}} \log^{\frac{3}{2}+\delta} n), & \zeta \geq 1, \quad 0 < \min\{\psi - 1, \gamma\} < 1, \\ O(n^{\frac{\zeta}{\psi}-\zeta} + n^{-\frac{1}{2}} \log^{\frac{3}{2}+\delta} n), & 0 < \zeta < 1, \quad \min\{\psi - 1, \gamma\} > 1, \\ O(n^{\frac{\zeta}{\psi}-\zeta} + n^{-\frac{1}{2}+\delta}), & 0 < \zeta < 1, \quad \min\{\psi - 1, \gamma\} = 1, \\ O(n^{\frac{\zeta}{\psi}-\zeta} + n^{-\frac{\min\{\psi-1,\gamma\}}{2}} \log^{\frac{3}{2}+\delta} n), & 0 < \zeta < 1, \quad 0 < \min\{\psi - 1, \gamma\} < 1. \end{cases}$$

Note that the previous theorem combined with a decaying bound on $d(\bar{W}_n, \sigma_n Z)$ yields quenched CLTs. Examples of quenched CLTs derived by applying the previous theorem are given in the last two subsections of this section.

Next we will consider the question: when $\sigma^2 > 0$? For this purpose define

$$\varphi^{(2)}(n, \omega)(x, y) = (\varphi(n, \omega)x, \varphi(n, \omega)y)$$

on the product space $X \times X$. Define the projections $\Pi_1(\omega, x, y) = \omega$, $\Pi_2(\omega, x, y) = x$ and $\Pi_3(\omega, x, y) = y$ on $\Omega \times X \times X$. The assumption below is needed in the lemma that considers the conditions for $\sigma^2 > 0$.

Standing Assumption (SA5). Assume there exists an invariant measure $\mathbf{P}^{(2)}$ for the RDS $\varphi^{(2)}$ that is symmetric in the sense that

$$\int h(\omega, x, y) \, d\mathbf{P}^{(2)}(\omega, x, y) = \int h(\omega, y, x) \, d\mathbf{P}^{(2)}(\omega, x, y) \tag{28}$$

for all bounded measurable $h : \Omega \times X \times X \to \mathbb{R}$. The common marginal

$$\mathbf{P} = (\Pi_1 \times \Pi_2)_* \mathbf{P}^{(2)} = (\Pi_1 \times \Pi_3)_* \mathbf{P}^{(2)} \tag{29}$$

is then trivially an invariant measure for the RDS φ. Moreover, assume

$$\lim_{i \to \infty} \bar{\mathbb{E}}[\mu(f_i)] = \int f(x) \, d\mathbf{P}(\omega, x), \tag{30}$$

$$\lim_{i \to \infty} \bar{\mathbb{E}}[\mu(f_i f_{i+k})] = \int f(x) \, f(\varphi(k, \omega, x)) \, d\mathbf{P}(\omega, x) \tag{31}$$

and

$$\lim_{i \to \infty} \bar{\mathbb{E}}[\mu(f_i)\mu(f_{i+k})] = \int f(x) \, f(\varphi(k, \omega, y)) \, d\mathbf{P}^{(2)}(\omega, x, y) \tag{32}$$

are satisfied. ■

Define the sequence of functions $Z_n \colon \Omega \times X \times X \to \mathbb{R}$ by

$$Z_n(\omega, x, y) = S_n(\omega, x) - S_n(\omega, y),$$

where $S_n = \sqrt{n} W_n$. Define the skew-product $\Phi^{(2)} \colon \Omega \times X \times X \to \Omega \times X \times X$ by

$$\Phi^{(2)}(\omega, x, y) = (\tau\omega, \varphi^{(2)}(1, \omega)(x, y)).$$

Set

$$F(\omega, x, y) = f(x) - f(y).$$

In the lemma below it is assumed that conditions (SA1) and (SA3)–(SA5) are satisfied. Note that the strong-mixing assumption (SA2) is unnecessary here.

Lemma 7.3. *Suppose $\eta(n) = Cn^{-\psi}$, $\psi > 2$.*

(i) $\sigma^2 = 0$ is equivalent to each of the following conditions:

(a) $\sup_{n\geq 0}\int Z_n^2\,\mathrm{d}\mathbf{P}^{(2)} < \infty$.
(b) *There exists* $G \in L^2(\mathbf{P}^{(2)})$ *such that* $F = G - G\circ\Phi^{(2)}$.

(ii) $\sigma^2 > 0$ *is equivalent to each of the following conditions:*
(a) $\sup_{n\geq 0}\int Z_n^2\,\mathrm{d}\mathbf{P}^{(2)} = \infty$.
(b) *There exist* $c > 0$ *and* $N > 0$ *such that* $\int Z_n^2\,\mathrm{d}\mathbf{P}^{(2)} \geq cn$ *for all* $n \geq N$.

(iii) *If* $\zeta > 1$, *then* $\sigma^2 > 0$ *is equivalent to each of the following conditions:*
(a) $\sup_{n\geq 1} n^{-\frac{1}{\psi}}\int Z_n^2\,\mathrm{d}(\mathbb{P}\otimes\mu\otimes\mu) = \infty$.
(b) $\sup_{n\geq 1} n^{-\frac{1}{\psi}}\,\mathbb{E}\,\mathrm{Var}_\mu(S_n) = \infty$.
(c) *There exist* $c > 0$ *and* $N > 0$ *such that* $\int Z_n^2\,\mathrm{d}(\mathbb{P}\otimes\mu\otimes\mu) \geq cn$ *for all* $n \geq N$.
(d) *There exist* $c > 0$ *and* $N > 0$ *such that* $\mathbb{E}\,\mathrm{Var}_\mu(S_n) \geq cn$ *for all* $n \geq N$.

(iv) *If* $\mathbb{P}$ *is stationary, then* $\sigma^2 = 0$ *is equivalent to each of the following conditions:*
(a) $\sup_{n\geq 1}\int Z_n^2\,\mathrm{d}(\mathbb{P}\otimes\mu\otimes\mu) < \infty$.
(b) $\sup_{n\geq 1}\mathbb{E}\,\mathrm{Var}_\mu(S_n) < \infty$.

(v) *If* $\mathbb{P}$ *is stationary, then* $\sigma^2 > 0$ *is equivalent to each of the following conditions:*
(a) $\sup_{n\geq 1}\int Z_n^2\,\mathrm{d}(\mathbb{P}\otimes\mu\otimes\mu) = \infty$.
(b) $\sup_{n\geq 1}\mathbb{E}\,\mathrm{Var}_\mu(S_n) = \infty$.
(c) *There exist* $c > 0$ *and* $N > 0$ *such that* $\int Z_n^2\,\mathrm{d}(\mathbb{P}\otimes\mu\otimes\mu) \geq cn$ *for all* $n \geq N$.
(d) *There exist* $c > 0$ *and* $N > 0$ *such that* $\mathbb{E}\,\mathrm{Var}_\mu(S_n) \geq cn$ *for all* $n \geq N$.

7.3. **Random expanding circle maps.** In this subsection we study an RDS where the transformations T_{ω_0} are expanding circle maps picked at random from the set $\mathcal{M}$ defined in Subsection 5.4. We study the limit variance of $W = W(N) = \sum_{i=0}^{N-1} f\circ T_{\omega_i}\circ\cdots\circ T_{\omega_1} - \mu(f\circ T_{\omega_i}\circ\cdots\circ T_{\omega_1})$[7].

To be more precise, as in the previous subsection each index ω_i is drawn randomly from a probability space $(\Omega,\mathcal{F},\mathbb{P}) = (\Omega_0^{\mathbb{N}},\mathcal{E}^{\mathbb{N}},\mathbb{P})$, where $(\Omega_0,\mathcal{E})$ is a measurable space. We assume the following about the random dynamical system in question:

Assumption (RDS)
i) Each $T_{\omega_i}\in\mathcal{M}$.
ii) The law $\mathbb{P}$ is stationary, i.e., the shift $\tau:\Omega\to\Omega : (\tau(\omega))_i = \omega_{i+1}$ preserves $\mathbb{P}$.
iii) The random selection process is strong mixing satisfying

$$\sup_{i\geq 1}\ \sup_{A\in\mathcal{F}_1^i,\,B\in\mathcal{F}_{i+n}^\infty}|\mathbb{P}(AB)-\mathbb{P}(A)\,\mathbb{P}(B)| \leq Cn^{-\gamma}$$

for each $n\geq 1$, where $\gamma > 0$.
iv) The map

$$(\omega,x)\mapsto T_{\omega_n}\circ\cdots\circ T_{\omega_1}(x)$$

is measurable from $\mathcal{F}\otimes\mathcal{B}$ to $\mathcal{B}$ for every $n\in\mathbb{N}_0$.

Define $\sigma_N^2(\omega) = \sigma_N^2 = \mathrm{Var}_\mu W(N)$ and $\sigma^2 = \lim_{N\to\infty}\mathbb{E}\sigma_N^2$, when the limit exists. The next theorem from article **(B)** gives a quenched CLT with a rate of convergence that holds for almost every sequence of transformations.

Theorem 7.4. *Assume that (RDS) is satisfied. Then* $\sigma > 0$ *if and only if*

$$\sup_{N\geq 1} N\,\mathbb{E}\mu(W^2) = \infty.$$

[7]If we were using the notations in the previous subsection, $W(N)$ would have been denoted by $\bar{W}_N$

Furthermore if $\sigma > 0$ holds, then for arbitrary $\delta > 0$ and almost every ω

$$d_{\mathscr{W}}(W(N), \sigma Z) = \begin{cases} O(N^{-\frac{1}{2}} \log^{\frac{3}{2}+\delta} N), & \gamma > 1, \\ O(N^{-\frac{1}{2}+\delta}), & \gamma = 1, \\ O(N^{-\frac{\gamma}{2}} \log^{\frac{3}{2}+\delta} N), & 0 < \gamma < 1, \end{cases}$$

where $\sigma^2 = \sum_{k=0}^{\infty} (2 - \delta_{k0}) \lim_{i\to\infty} \mathbb{E}[\mu(f_i f_{i+k}) - \mu(f_i)\mu(f_{i+k})]$.

7.4. **Random Pomeau-Manneville Maps.** The last theorem given in this introduction is for an RDS of Pomeau-Manneville maps. This theorem has a lot of similarity to the one above.

Here $(T_{\omega_i})_{i=1}^{\infty}$ is a random sequence of transformations on $\mathcal{P}$ such that each $(\omega_i)_{i=1}^{\infty}$ is drawn randomly from a probability space $(\Omega, \mathcal{F}, \mathbb{P}) = ([0, \beta_*]^{\mathbb{N}}, \mathcal{E}^{\mathbb{N}}, \mathbb{P})$. Here $\mathcal{E}$ is the Borel algebra of $[0, \beta_*]$. P-M maps T_{ω_i} are defined as in (15).

For this system of random Pomeau-Manneville maps we define in article **(C)** a set of conditions, which is also abbreviated by (RDS). It is the same as the the set of conditions for random expanding circle maps in the previous subsection, with only exceptions being that $T_{\omega_i} \in \mathcal{P}$ instead of $\mathcal{M}$ and that condition (iv) is not explicitly required since it follows from the properties of the model.

Theorem 7.5. *Assume that (RDS) is satisfied with $\beta_* < 1/3$. Then*

$$\sigma^2 = \sum_{k=0}^{\infty} (2 - \delta_{k0}) \lim_{i\to\infty} \mathbb{E}[\mu(f_i f_{i+k}) - \mu(f_i)\mu(f_{i+k})]$$

is well-defined and non-negative. We have $\sigma > 0$ if and only if

$$\sup_{N\geq 1} N\, \mathbb{E}\mu(W^2) = \infty.$$

Furthermore if $\sigma > 0$ holds, then for arbitrary $\delta > 0$ and almost every ω

$$d_{\mathscr{W}}(W(N), \sigma Z) = \begin{cases} O(N^{\beta_* - \frac{1}{2}} (\log N)^{\frac{1}{\beta_*}}), & \gamma \geq 1, \\ O(N^{\beta_* - \frac{1}{2}} (\log N)^{\frac{1}{\beta_*}}) + O(N^{-\frac{\gamma}{2}} (\log N)^{\frac{3}{2}+\delta}), & 0 < \gamma < 1. \end{cases}$$

References

[1] Mohamed Abdelkader and Romain Aimino. On the quenched central limit theorem for random dynamical systems. *J. Phys. A*, 49(24):244002, 13, 2016. doi:10.1088/1751-8113/49/24/244002.

[2] Romain Aimino, Huyi Hu, Matthew Nicol, Andrei Török, and Sandro Vaienti. Polynomial loss of memory for maps of the interval with a neutral fixed point. *Discrete Contin. Dyn. Syst.*, 35(3):793–806, 2015. doi:10.3934/dcds.2015.35.793.

[3] Romain Aimino, Matthew Nicol, and Sandro Vaienti. Annealed and quenched limit theorems for random expanding dynamical systems. *Probab. Theory Related Fields*, 162(1-2):233–274, 2015. doi:10.1007/s00440-014-0571-y.

[4] Romain Aimino and Jérôme Rousseau. Concentration inequalities for sequential dynamical systems of the unit interval. *Ergodic Theory Dynam. Systems*, 36(8):2384–2407, 2016. doi:10.1017/etds.2015.19.

[5] Marios Antoniou and Ian Melbourne. Rate of convergence in the weak invariance principle for deterministic systems. *Comm. Math. Phys.*, 369(3):1147–1165, 2019. doi:10.1007/s00220-019-03334-6.

[6] Ludwig Arnold. *Random dynamical systems*. Springer Monographs in Mathematics. Springer-Verlag, Berlin, 1998. doi:10.1007/978-3-662-12878-7.

[7] Richard Arratia, Larry Goldstein, and Louis Gordon. Poisson approximation and the Chen-Stein method. *Statist. Sci.*, 5(4):403–434, 1990. With comments and a rejoinder by the authors. URL: http://links.jstor.org/sici?sici=0883-4237(199011)5:4<403:PAATCM>2.0.CO;2-6&origin=MSN.

[8] Arvind Ayyer, Carlangelo Liverani, and Mikko Stenlund. Quenched CLT for random toral automorphism. *Discrete Contin. Dyn. Syst.*, 24(2):331–348, 2009. doi:10.3934/dcds.2009.24.331.

[9] Wael Bahsoun and Christopher Bose. Corrigendum: Mixing rates and limit theorems for random intermittent maps (2016 *nonlinearity* 29 1417) [MR3476513]. *Nonlinearity*, 29(12):C4, 2016. doi:10.1088/0951-7715/29/12/C4.

[10] Wael Bahsoun, Christopher Bose, and Yuejiao Duan. Decay of correlation for random intermittent maps. *Nonlinearity*, 27(7):1543–1554, 2014. doi:10.1088/0951-7715/27/7/1543.

[11] Wael Bahsoun, Christopher Bose, and Marks Ruziboev. Quenched decay of correlations for slowly mixing systems. 2017. Preprint. arXiv:1706.04158.

[12] B. Berckmoes, R. Lowen, and J. Van Casteren. Stein's method and a quantitative Lindeberg CLT for the Fourier transforms of random vectors. *J. Math. Anal. Appl.*, 433(2):1441–1458, 2016. doi:10.1016/j.jmaa.2015.08.040.

[13] Sourav Chatterjee and Elizabeth Meckes. Multivariate normal approximation using exchangeable pairs. *ALEA Lat. Am. J. Probab. Math. Stat.*, 4:257–283, 2008.

[14] Jianyu Chen, Yun Yang, and Hong-Kun Zhang. Non-stationary almost sure invariance principle for hyperbolic systems with singularities. *Journal of Statistical Physics*, 2018. doi:10.1007/s10955-018-2107-9.

[15] Louis H. Y. Chen. Poisson approximation for dependent trials. *Ann. Probability*, 3(3):534–545, 1975. doi:10.1214/aop/1176996359.

[16] Louis H. Y. Chen, Larry Goldstein, and Qi-Man Shao. *Normal approximation by Stein's method*. Probability and its Applications (New York). Springer, Heidelberg, 2011. doi:10.1007/978-3-642-15007-4.

[17] Jean-Pierre Conze and Albert Raugi. Limit theorems for sequential expanding dynamical systems on $[0,1]$. In *Ergodic theory and related fields*, volume 430 of *Contemp. Math.*, pages 89–121. Amer. Math. Soc., Providence, RI, 2007. doi:10.1090/conm/430/08253.

[18] Manfred Denker, Mikhail Gordin, and Anastasya Sharova. A Poisson limit theorem for toral automorphisms. *Illinois J. Math.*, 48(1):1–20, 2004. URL: http://projecteuclid.org/euclid.ijm/1258136170.

[19] Neil Dobbs and Mikko Stenlund. Quasistatic dynamical systems. *Ergodic Theory and Dynamical Systems*, 2016. doi:10.1017/etds.2016.9.

[20] D. Dragičević, G. Froyland, C. González-Tokman, and S. Vaienti. Almost sure invariance principle for random piecewise expanding maps. *Nonlinearity*, 31(5):2252–2280, 2018. doi:10.1088/1361-6544/aaaf4b.

[21] D. Dragičević, G. Froyland, C. González-Tokman, and S. Vaienti. A spectral approach for quenched limit theorems for random expanding dynamical systems. *Comm. Math. Phys.*, 360(3):1121–1187, 2018. doi:10.1007/s00220-017-3083-7.

[22] Loïc Dubois. An explicit Berry-Esséen bound for uniformly expanding maps on the interval. *Israel J. Math.*, 186:221–250, 2011. doi:10.1007/s11856-011-0137-y.

[23] Rick Durrett. *Probability: theory and examples*, volume 31 of *Cambridge Series in Statistical and Probabilistic Mathematics*. Cambridge University Press, Cambridge, fourth edition, 2010. doi:10.1017/CBO9780511779398.

[24] Werner Ehm. Binomial approximation to the Poisson binomial distribution. *Statist. Probab. Lett.*, 11(1):7–16, 1991. doi:10.1016/0167-7152(91)90170-V.

[25] Hans Fischer. *A history of the central limit theorem*. Sources and Studies in the History of Mathematics and Physical Sciences. Springer, New York, 2011. From classical to modern probability theory. doi:10.1007/978-0-387-87857-7.

[26] Ana Cristina Freitas, Jorge Freitas, and Sandro Vaienti. Extreme value laws for sequences of intermittent maps. *Proceedings of the American Mathematical Society*, 2018. doi:10.1090/proc/13892.

[27] Ana Cristina Moreira Freitas, Jorge Milhazes Freitas, and Sandro Vaienti. Extreme value laws for non stationary processes generated by sequential and random dynamical systems. *Ann. Inst. H. Poincaré Probab. Statist.*, 53(3):1341–1370, 08 2017. doi:10.1214/16-AIHP757.

[28] Jason Fulman and Nathan Ross. Exponential approximation and Stein's method of exchangeable pairs. *ALEA Lat. Am. J. Probab. Math. Stat.*, 10(1):1–13, 2013.

[29] Robert E. Gaunt. Rates of convergence in normal approximation under moment conditions via new bounds on solutions of the Stein equation. *J. Theoret. Probab.*, 29(1):231–247, 2016. doi:10.1007/s10959-014-0562-z.

[30] Larry Goldstein and Yosef Rinott. Multivariate normal approximations by stein's method and size bias couplings. *Journal of Applied Probability*, 33(1):1–17, 1996. doi:10.2307/3215259.

[31] Mikhail Gordin. A homoclinic version of the central limit theorem. *Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI)*, 184(Issled. po Mat. Statist. 9):80–91, 321, 1990. doi:10.1007/BF01254269.

[32] Mikhail Gordin and Manfred Denker. The Poisson limit for automorphisms of two-dimensional tori defined by continued fractions. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)*, 408(Veroyatnost i Statistika. 18):131–153, 325–326, 2012.

[33] Sébastien Gouëzel. Berry-Esseen theorem and local limit theorem for non uniformly expanding maps. *Ann. Inst. H. Poincaré Probab. Statist.*, 41(6):997–1024, 2005. doi:10.1016/j.anihpb.2004.09.002.

[34] Friedrich Götze. On the rate of convergence in the multivariate clt. *The Annals of Probability*, pages 724–739, 1991. doi:10.1214/aop/1176990448.

[35] N. T. A. Haydn. Entry and return times distribution. *Dyn. Syst.*, 28(3):333–353, 2013. doi:10.1080/14689367.2013.822459.

[36] Nicolai Haydn, Matthew Nicol, Andrew Török, and Sandro Vaienti. Almost sure invariance principle for sequential and non-stationary dynamical systems. *Transactions of the American Mathematical Society*, 2017. doi:10.1090/tran/6812.

[37] Nicolai Haydn and Fan Yang. Entry times distribution for mixing systems. *J. Stat. Phys.*, 163(2):374–392, 2016. doi:10.1007/s10955-016-1487-y.

[38] Christophe Jan. Vitesse de convergence dans le TCL pour des chaînes de Markov et certains processus associés à des systèmes dynamiques. *C. R. Acad. Sci. Paris Sér. I Math.*, 331(5):395–398, 2000. doi:10.1016/S0764-4442(00)01615-3.

[39] Anatole Katok and Boris Hasselblatt. *Introduction to the modern theory of dynamical systems*, volume 54 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1995. With a supplementary chapter by Katok and Leonardo Mendoza. doi:10.1017/CBO9780511809187.

[40] Christoph Kawan. Expanding and expansive time-dependent dynamics. *Nonlinearity*, 28(3):669–695, 2015. doi:10.1088/0951-7715/28/3/669.

[41] Yuri Kifer. Limit theorems for random transformations and processes in random environments. *Trans. Amer. Math. Soc.*, 350(4):1481–1518, 1998. doi:10.1090/S0002-9947-98-02068-6.

[42] Jonathan L. King. On M. Gordin's homoclinic question. *Internat. Math. Res. Notices*, (5):203–212, 1997. doi:10.1155/S1073792897000147.

[43] A Korepanov, Z Kosloff, and I Melbourne. Martingale–coboundary decomposition for families of dynamical systems. In *Annales de l'Institut Henri Poincare (C) Non Linear Analysis*. Elsevier, 2017. doi:10.1016/j.anihpc.2017.08.005.

[44] A. Lasota and James A. Yorke. On the existence of invariant measures for piecewise monotonic transformations. *Trans. Amer. Math. Soc.*, 186:481–488 (1974), 1973. doi:10.2307/1996575.

[45] Juho Leppänen. Intermittent quasistatic dynamical systems: weak convergence of fluctuations. *Nonauton. Dyn. Syst.*, 5:8–34, 2018. doi:10.1515/msds-2018-0002.

[46] Juho Leppänen and Mikko Stenlund. Quasistatic dynamics with intermittency. *Math. Phys. Anal. Geom.*, 19(2):Art. 8, 23, 2016. doi:10.1007/s11040-016-9212-2.

[47] Juho Leppänen. Functional correlation decay and multivariate normal approximation for non-uniformly expanding maps. *Nonlinearity*, 30(11):4239, 2017. URL: http://stacks.iop.org/0951-7715/30/i=11/a=4239.

[48] Carlangelo Liverani. Central limit theorem for deterministic systems. In *International Conference on Dynamical Systems (Montevideo, 1995)*, volume 362 of *Pitman Res. Notes Math. Ser.*, pages 56–75. Longman, Harlow, 1996.

[49] Carlangelo Liverani, Benoît Saussol, and Sandro Vaienti. A probabilistic approach to intermittency. *Ergodic Theory Dynam. Systems*, 19(3):671–685, 1999. URL: http://dx.doi.org/10.1017/S0143385799133856.

[50] Ho Ming Luk. *Stein's method for the Gamma distribution and related statistical applications*. ProQuest LLC, Ann Arbor, MI, 1994. Thesis (Ph.D.)–University of Southern California. URL: http://gateway.proquest.com/openurl?url_ver=Z39.88-2004&rft_val_fmt=info:ofi/fmt:kev:mtx:dissertation&res_dat=xri:pqdiss&rft_dat=xri:pqdiss:9601018.

[51] Elizabeth Meckes. On Stein's method for multivariate normal approximation. In *High dimensional probability V: the Luminy volume*, volume 5 of *Inst. Math. Stat. Collect.*, pages 153–178. Inst. Math. Statist., Beachwood, OH, 2009. doi:10.1214/09-IMSCOLL511.

[52] Elizabeth Meckes. Approximation of projections of random vectors. *J. Theoret. Probab.*, 25(2):333–352, 2012. doi:10.1007/s10959-010-0299-2.

[53] Anushaya Mohapatra and William Ott. Memory loss for nonequilibrium open dynamical systems. *Discrete Contin. Dyn. Syst.*, 34(9):3747–3759, 2014. doi:10.3934/dcds.2014.34.3747.

[54] Péter Nándori, Domokos Szász, and Tamás Varjú. A Central Limit Theorem for Time-Dependent Dynamical Systems. *Journal of Statistical Physics*, 146(6):1213–1220, MAR 2012. doi:10.1007/s10955-012-0451-8.

[55] Matthew Nicol, Andrew Török, and Sandro Vaienti. Central limit theorems for sequential and random intermittent dynamical systems. *Ergodic Theory Dynam. Systems*, 38(3):1127–1153, 2018. doi:10.1017/etds.2016.69.

[56] Ivan Nourdin, Giovanni Peccati, and Anthony Réveillac. Multivariate normal approximation using Stein's method and Malliavin calculus. *Ann. Inst. Henri Poincaré Probab. Stat.*, 46(1):45–58, 2010. doi:10.1214/08-AIHP308.

[57] William Ott, Mikko Stenlund, and Lai-Sang Young. Memory loss for time-dependent dynamical systems. *Mathematical Research Letters*, 16(3):463–475, 2009.

[58] Françoise Pène. Rates of convergence in the CLT for two-dimensional dispersive billiards. *Comm. Math. Phys.*, 225(1):91–119, 2002. doi:10.1007/s002201000573.

[59] Françoise Pène. Multiple decorrelation and rate of convergence in multidimensional limit theorems for the Prokhorov metric. *Ann. Probab.*, 32(3B):2477–2525, 2004. doi:10.1214/009117904000000036.

[60] Françoise Pène. Rate of convergence in the multidimensional central limit theorem for stationary processes. Application to the Knudsen gas and to the Sinai billiard. *Ann. Appl. Probab.*, 15(4):2331–2392, 2005. doi:10.1214/105051605000000476.

[61] Mark Pollicott and Richard Sharp. Invariance principles for interval maps with an indifferent fixed point. *Comm. Math. Phys.*, 229(2):337–346, 2002. doi:10.1007/s00220-002-0685-4.

[62] Yiannis Psiloyenis. Mixing conditions and return times on markov towers, 2008. Dissertation, University of Southern California. URL: http://search.proquest.com/docview/304461750?accountid=11365.

[63] M. Ratner. The central limit theorem for geodesic flows on n-dimensional manifolds of negative curvature. *Israel J. Math.*, 16:181–197, 1973. doi:10.1007/BF02757869.

[64] Gesine Reinert and Adrian Röllin. Multivariate normal approximation with Stein's method of exchangeable pairs under a general linearity condition. *Ann. Probab.*, 37(6):2150–2173, 2009. doi:10.1214/09-AOP467.

[65] Yosef Rinott and Vladimir Rotar. A multivariate CLT for local dependence with $n^{-1/2}\log n$ rate and applications to multivariate graph related statistics. *J. Multivariate Anal.*, 56(2):333–350, 1996. doi:10.1006/jmva.1996.0017.

[66] Adrian Röllin. Stein's method in high dimensions with applications. *Ann. Inst. Henri Poincaré Probab. Stat.*, 49(2):529–549, 2013.

[67] Ja. G. Sinaĭ. The central limit theorem for geodesic flows on manifolds of constant negative curvature. *Soviet Math. Dokl.*, 1:983–987, 1960.

[68] Charles Stein. A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. In *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability (Univ. California, Berkeley, Calif., 1970/1971), Vol. II: Probability theory*, pages 583–602. Univ. California Press, Berkeley, Calif., 1972.

[69] Mikko Stenlund. Non-stationary compositions of Anosov diffeomorphisms. *Nonlinearity*, 24:2991–3018, 2011. doi:10.1088/0951-7715/24/10/016.

[70] Mikko Stenlund. An almost sure ergodic theorem for quasistatic dynamical systems. *Math. Phys. Anal. Geom.*, 19(3):Art. 14, 18, 2016. doi:10.1007/s11040-016-9217-x.

[71] Yaofeng Su. Vector-valued almost sure invariance principle for non-stationary dynamical systems. 2019. Preprint. arXiv:1903.09763.